

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:10:21 ; Search time 27 Seconds
(without alignments)
2231.389 Million cell updates/sec

Title: US-10-089-678-1

Perfect score: 6044
Sequence: 1 MSPNNQNEYILDASSSTSV.....MSETGVNIESVELVLEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/6C_COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4007	66.3	1157	US-07-876-280-30	Sequence 30, Appl
2	4007	66.3	1157	US-07-812-180A-2	Sequence 2, Appl
3	4007	66.3	1157	US-08-315-468-2	Sequence 2, Appl
4	4007	66.3	1157	US-07-941-650A-2	Sequence 2, Appl
5	3662.5	60.6	1149	US-07-915-203-2	Sequence 2, Appl
6	3662.5	60.6	1149	US-08-272-887-2	Sequence 2, Appl
7	3662.5	60.6	1149	US-08-789-449-2	Sequence 2, Appl
8	3152	52.2	1169	US-08-315-468-4	Sequence 4, Appl
9	3123	51.7	1157	US-08-532-547-5	Sequence 5, Appl
10	3123	51.7	1157	US-08-379-656B-5	Sequence 5, Appl
11	3123	51.7	1157	US-08-455-838-5	Sequence 5, Appl
12	3123	51.7	1157	US-09-019-809-5	Sequence 5, Appl
13	3123	51.7	1157	US-09-471-177-5	Sequence 5, Appl
14	3100.5	51.3	1156	US-09-002-285-72	Sequence 72, Appl
15	3100.5	51.3	1156	US-08-589-477-72	Sequence 72, Appl
16	3100.5	51.3	1156	US-09-661-322A-28	Sequence 28, Appl
17	2825.5	46.7	1150	US-09-002-285-74	Sequence 74, Appl
18	2825.5	46.7	1150	US-09-589-477-74	Sequence 74, Appl
19	2715	44.9	1169	US-08-542-921-2	Sequence 2, Appl
20	2715	44.9	1169	US-08-880-685-2	Sequence 2, Appl
21	2715	44.9	1169	US-08-880-684-2	Sequence 2, Appl
22	2710	44.8	1134	US-09-002-285-76	Sequence 76, Appl
23	2710	44.8	1134	US-09-589-477-76	Sequence 76, Appl
24	2549	42.2	1229	US-08-100-709-4	Sequence 4, Appl
25	2549	42.2	1229	US-08-176-865-4	Sequence 4, Appl
26	2549	42.2	1229	US-08-474-038-4	Sequence 4, Appl
27	2549	42.2	1229	US-08-779-046-4	Sequence 4, Appl

28 2549 42.2 1229 2 US-08-881-340-4 Sequence 4, Appl
29 2525 41.8 1227 3 US-09-053-549-2 Sequence 2, Appl
30 2519.5 41.7 1207 1 US-07-951-715A-7 Sequence 7, Appl
31 2519.5 41.7 1207 2 US-08-459-448A-7 Sequence 7, Appl
32 2519.5 41.7 1207 3 US-08-459-595A-7 Sequence 7, Appl
33 2519.5 41.7 1207 3 US-08-459-504B-7 Sequence 7, Appl
34 2519.5 41.7 1207 3 US-08-459-444-7 Sequence 7, Appl
35 2519.5 41.7 1207 3 US-09-053-549-8 Sequence 8, Appl
36 2519.5 41.7 1207 4 US-09-547-422-7 Sequence 7, Appl
37 2459 40.7 1227 1 US-08-448-170-8 Sequence 8, Appl
38 2459 40.7 1227 3 US-08-961-803-9 Sequence 9, Appl
39 2458.5 40.7 1138 1 US-07-973-320-2 Sequence 2, Appl
40 2457 40.7 1227 4 US-09-661-322A-63 Sequence 63, Appl
41 2454.5 40.6 1186 3 US-09-178-252-23 Sequence 23, Appl
42 2454.5 40.6 1186 4 US-09-828-660-23 Sequence 23, Appl
43 2452.5 40.6 1228 1 US-09-661-322A-38 Sequence 38, Appl
44 2441.5 40.4 1138 1 US-07-973-320-4 Sequence 4, Appl
45 2237.5 37.0 1177 3 US-08-754-490-14 Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-07-876-280-30
; Sequence 30, Application US/07876280
; Patent No. 5262158
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Bagley, Angela L.
; TITLE OF INVENTION: No. 5262158el Bacillus thuringiensis Isolates for
; TITLE OF INVENTION: Controlling Acarides
; NUMBER OF INVENTIONS: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/876,280
; FILING DATE: 19920430
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/S 104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: PS50C
; IMMEDIATE SOURCE:
; CLONE: E. coli NM522 (pMYC2320) NRRL B-18769
US-07-876-280-30

Query Match 66.3%; Score 4007; DB 1; Length 1157;
Best Local Similarity 66.1%; Pred. No. 2.3e-285;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGEGNPELFGNP 60
DB 1 MSPNNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGEGNPELFGNP 60

QY 61 ETPFISSTVQTGIGIVGVVAGLGVFPAGQIASFYFIVQGLWPSSTVSWENIMKQVED 120
DB 61 ETPFISSTVQTGIGIVGVVAGLGVFPAGQIASFYFIVQGLWPSSTVSWENIMKQVED 120

QY 121 LIDOKITDVRKTAALAGLQGLGVVQKSLKWLKLNENRNDTPARSVVVQYIYALBDFV 180
DB 121 LIDOKITDVRKTAALAGLQGLGVVQKSLKWLKLNENRNDTPARSVVVQYIYALBDFV 180

QY 181 AKIPSPAISQOEYPLSVVAAQANLHLLLRDASIFCAEWGFTGPGISFYFDRQVTRTAQ 240
DB 181 AKIPSPAISQOEYPLSVVAAQANLHLLLRDASIFCAEWGFTGPGISFYFDRQVTRTAQ 240

QY 241 YSDYCVKWNVTGLDKLGTNAASWLKYHQFRREMTLLVLDLVALFNYDTRTYPIETTAQ 300
DB 241 YSDYCVKWNVTGLDKLGTNAASWLKYHQFRREMTLLVLDLVALFNYDTRTYPIETTAQ 300

QY 301 LTRVNVTDPIVFNRETSGGCRWSLNSDISFSEVESAVIRSHPLFDLISELIEFYTRAG 360
DB 301 LTRVNVTDPIVFNRETSGGCRWSLNSDISFSEVESAVIRSHPLFDLISELIEFYTRAG 360

QY 361 LPLNNTYEYLYWYGHISKYKNT--NASSALERNYGTITTSNKIKYDILANKDIFQVRSIGAD 419
DB 361 LPLNNTYEYLYWYGHISKYKNT--NASSALERNYGTITTSNKIKYDILANKDIFQVRSIGAD 419

QY 420 LANYAQVGVVYASFTLLDKNTGSGVGGFTYKSPHTTMQVCTQNYNTIDEIPENE-P 478
DB 420 LANYAQVGVVYASFTLLDKNTGSGVGGFTYKSPHTTMQVCTQNYNTIDEIPENE-P 478

QY 479 LSGYSHRLSHITSYFSGKNASSPARYGNLPVANTHRGADVTNTVYSKTIQIPVKAH 538
DB 479 LSGYSHRLSHITSYFSGKNASSPARYGNLPVANTHRGADVTNTVYSKTIQIPVKAH 538

QY 539 TLVSGTIVKPGFTGNILKRTSSGPLAYTSVKSPLSORVRAIRVASTTNLRLFT 598
DB 539 TLVSGTIVKPGFTGNILKRTSSGPLAYTSVKSPLSORVRAIRVASTTNLRLFT 598

QY 599 ISGTRYSINNVKTMKGGDLTNTFDLATIGTAPTSYNSDSLTVGADSPAGSGVYVD 658
DB 599 ISGTRYSINNVKTMKGGDLTNTFDLATIGTAPTSYNSDSLTVGADSPAGSGVYVD 658

QY 659 KPELIPVNATFEAEEDLVAKKAVNGLFTSKDALQTSVTDYQVNOAANLVECLSDLYP 718
DB 659 KPELIPVNATFEAEEDLVAKKAVNGLFTSKDALQTSVTDYQVNOAANLVECLSDLYP 718

QY 719 NEKRLMDVAKEAKRLVQARNLLQDTGFNRINGENGWGTSGTIEVAGDVLFKDRSLRLT 778
DB 719 NEKRLMDVAKEAKRLVQARNLLQDTGFNRINGENGWGTSGTIEVAGDVLFKDRSLRLT 778

QY 779 SAREIDTPTLYQOQIDSLKPYTRYKLGFGSSQDLKILRHRANOIVKXVPDN 838
DB 779 SAREIDTPTLYQOQIDSLKPYTRYKLGFGSSQDLKILRHRANOIVKXVPDN 838

QY 839 LLPDVLPVNSCGGIDRCSEQOQYDANLLENNGENGMWSSDASHAFHIDTGTIDNENT 898
DB 839 LLPDVLPVNSCGGIDRCSEQOQYDANLLENNGENGMWSSDASHAFHIDTGTIDNENT 898

QY 899 GIWVFKIPTNGYATLGNLELVEEGLPSGETLERAQQOQEQMDKMKRGASEKAYYA 958
DB 899 GIWVFKIPTNGYATLGNLELVEEGLPSGETLERAQQOQEQMDKMKRGASEKAYYA 958

QY 959 AKQAIIDLFAQYQDQKLSGVMSMDLAAQNLVQSIPYVYNDALPEIPGNMNTSFTELTN 1018
DB 959 AKQAIIDLFAQYQDQKLSGVMSMDLAAQNLVQSIPYVYNDALPEIPGNMNTSFTELTN 1018

QY 1019 RLQQAANLYDRNAIPNGDFRNLGSDMNSATSDVNTQQLSDTSVLVIPNNNSOVSOQFTVQ 1078
DB 1019 RLQQAANLYDRNAIPNGDFRNLGSDMNSATSDVNTQQLSDTSVLVIPNNNSOVSOQFTVQ 1078

QY 1079 PNYRVYLVARTARKEGVGYVIRRGANQOTELTFTNICDDDTGVLSDADQTSYITKTVEFT 1138
DB 1079 PNYRVYLVARTARKEGVGYVIRRGANQOTELTFTNICDDDTGVLSDADQTSYITKTVEFT 1138

QY 1139 PSTEQWIDMSETEGVFNIESVELVLEE 1167
DB 1139 PSTEQWIDMSETEGVFNIESVELVLEE 1167

RESULT 2
US-07-812-180A-2
Sequence 2, Application US/07812180A
Patent No. 5366892
GENERAL INFORMATION:
APPLICANT: Poncerrada, Luis R
APPLICANT: Payne, Jewel M
TITLE OF INVENTION: No. 5366892el Coleopteran-Active Bacillus
TITLE OF INVENTION: thuringiensis Isolate and a No. 5366892el Gene Encoding a
TITLE OF INVENTION: Coleopteran-Active Toxin
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Roman Saliwanchik
STREET: 2421 N.W. 41st Street, Ste A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/812,180A
FILING DATE: 19920102
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, Roman
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1157 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: kumamotoensis
INDIVIDUAL ISOLATE: PS50C
IMMEDIATE SOURCE:
LIBRARY: Lambdagem (TM) - 11 LIBRARY OF LUIS
LIBRARY: FONCERRADA
CLONE: 50C
US-07-812-180A-2

Query Match 66.3%; Score 4007; DB 1; Length 1157;
Best Local Similarity 66.1%; Pred. No. 2.3e-285;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGEGNPELFGNP 60
DB 1 MSPNNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGEGNPELFGNP 60


```

Db 61 1129 PYTDMWIEISTEGTFYIESVELIVDVE 1157
Qy 121 LIDQKITDSVRKTALGLOGGLGVYQKSLKNWLENRNDTRARSVVTVQYIALDLDFV 180
Db 121 LVQKIEKYVKKALAEKGLGNALDVYQOSLEDWLENRNDATRSVSVNQFALDLNFV 180
Qy 181 AKIPSPAISQOEVLPLSVAAQANLHLLLRDASIFGAEGWFTPGIBSTFYDRQVTRTAQ 240
Db 181 SSIPSPAVSGHEVLLAVAAQVNLHLLLRDASIFGEWGFPGIBSTFYDRQVTRTAQ 240
Qy 241 YSDYCVKWTGDKLGTNAASWLVKXQPRRMTLLVLDLVALFNYDTRTPYPIETTAQ 300
Db 241 YSDYCVKWKYKIGDKLGTTSKSWNLVHQPRRMTLLVLDLVALFNYDTHMYPYPIETTAQ 300
Qy 301 LTRVYVTPVFNRETSGGFCRWLSNDSISFSEVESAVIRSPHLFDILSEIEFYTRAG 360
Db 301 LTRDVYDPIAFNIVTSTGFCNPSWTHSGILFVEVENNVIRPHLFDILLSVEINTSRG 360
Qy 361 LPLNTEYLEYVWCHSIKYKNT-NASALERNYGTITSNKIKYDILANKDIOFVRSLGAD 419
Db 361 ITLNDAYINYSGHTLKRRATDSTVYTYANTGRITSEK-NSFALEDRIEINSTVAN 419
Qy 420 LANYAQVYGPVYASFTLLDKNTGSGVGGFTYSKPHTTMVCQNTYNTIDEIPPENE-P 478
Db 420 LANYQAYGVPGSWHMKR--GTSTTAYLYSKTHALQGCQVIESSEDEIPDORTVP 477
Qy 479 LSRGYSHRLSHITSYSPKSNASSPARYGNLVPFANWTHRSADVNTVYSDKITQIPVVKAH 538
Db 478 VASYSHRLSHITSYSPKSNASSPARYGNLVPFANWTHRSADVNTVYSDKITQIPVVKAH 535
Qy 539 TLVSGTIVKPGFTGNILKRTSSGPLATSVSVKSPLSQRYRARIIRASTNLRFT 598
Db 536 MLYLGGSVVQPGFTGDKLRTNPSILGTFAVTVNGSLSQRYRARIIRASTNLRFT 594
Qy 599 ISGTRIVSINVNTMKNKDDLTWTFDLAGTFTFNSYSDSLTVGADSPASGGEVVD 658
Db 595 YLGDTEKRNFTKMGASLTETFKFASFTDFQPRETQDKILLSMGDFSGGQVYID 654
Qy 659 KFLIPVNAFPEABEDLDVAKKAVNGLTFSKKDALQTSVTDYVQVQANLVECLSDLYP 718
Db 655 RIEFIPVDETYEABQLEAAKAVNALFTNKDGLRPGVTDYEVNQANLVECLSDLYP 714
Qy 719 NEKRLMDVAKKRLVQARNLQDGFNRCINGENGTGTEVAGDVLFKDRLRLT 778
Db 715 NEKRLFDVAREAKRLSGARNLQDQDFQINGENGAASGTGTEVAGDVLFKDRLRLP 774
Qy 779 SAREIDTETVLYQOIDSILKPYTRYKLGFIGSSODLEIKLIRHRANOIVKXVNP 838
Db 775 GAREIDTETVLYQKVEGVLPTRYRLRGLFVSSQGLEIYTHQTNRIKXVNP 834
Qy 839 LLPVLPVNSCGGIDRCSEQQYVDANLALENNGENGMGSDSHAFSHIDTGBIDLNT 898
Db 835 LLPVLPVNSDGSINRCSEQQYVNSRL-----EGENRSGDAHEFSLPIDIGELDYNEA 888
Qy 899 GIWVVKIPTNGYATILGNLEVEBGLSETLERAOQOQOQWQKARKGASEKAYYA 958
Db 889 GIWVGKIPDTEGYATILGNLEVEBGLSETLERAOQOQOQWQKARKGASEKAYYA 948
Qy 959 AKQADRLPADYQDKLNSGVMSDMLAQNLSQISPYVYNDALPEIPGMNYSPTLNT 1018
Db 949 SKQAVDELYADYQDQQLNPVEITDLTAQDLQSIPIYVYVNEFPPEIPGMNYSPTLNT 1008
Qy 1019 RLQQANLYDLRNPAGDFRNLSDWNATSDVNVQQLSDTSVLVPIPNWNSQVSOQFTVQ 1078
Db 1009 RLQQANLYDORNAIPNGDFRNLSDWNATPGVEVQIINHTSVLVIPNWDQVSOQFTVQ 1068
Qy 1079 PNRYVLRVTKARKEGVDGVVIRIDGANOTETLTFNICDDDTGVLSDQTSYTKTVEFT 1138
Db 1069 PNORYVLRVTKARKEGVNGVYIRIDGNGQSETLTFNDSASDYDTNGVYNDQTYTKTVEFT 1128
Qy 1139 PSTEQWIDMSETGEVFNIESVELVLEE 1167

```

```

Db 1129 PYTDMWIEISTEGTFYIESVELIVDVE 1157
RESULT 4
US-07-941-650A-2
Sequence 2, Application US/07941650A
Patent No. 6294184
GENERAL INFORMATION:
APPLICANT: Uyeda, Kendrick A.
APPLICANT: Bradfisch, Gregory A.
TITLE OF INVENTION: Process for Controlling Lepidopteran Pests
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/941,650A
FILING DATE: 19920908
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/758,020
FILING DATE: 12-SEP-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/658,935
FILING DATE: 21-FEB-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/642,112
FILING DATE: 16-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1157 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: kumamotoensis
INDIVIDUAL ISOLATE: PS50C
IMMEDIATE SOURCE:
CLONE: E. coli NM522 (pMYC1638), NRRL B-18751
US-07-941-650A-2
Query Match 66.3%; Score 4007; DB 3; Length 1157;
Best Local Similarity 66.1%; Pred. No. 2.3e-285;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;
Qy 1 MSPNNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNMNYKDYLRMSEGNPELFGNP 60
Db 1 MSPNNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNMNYKDYLRMSEGNPELFGNP 60
Qy 61 ETFISSSTVOTGIGVQVGLGALGYPFAGQIASFYSFVIGQLWPSSTVSVWEMIKQVED 120
Db 61 ETFISSSTVOTGIGVQVGLGALGYPFAGQIASFYSFVIGQLWPSSTVSVWEMIKQVED 120
Qy 61 ETFISSSTVOTGIGVQVGLGALGYPFAGQIASFYSFVIGQLWPSSTVSVWEMIKQVED 120
Db 61 ETFISSSTVOTGIGVQVGLGALGYPFAGQIASFYSFVIGQLWPSSTVSVWEMIKQVED 120
Qy 121 LIDQKITDSVRKTALGLOGGLGVYQKSLKNWLENRNDTRARSVVTVQYIALDLDFV 180

```



```

Db 121 LVDQKLEKVKOKALAEKGLGALDVOYQOSEDMLNENRNDARTSRVSNQFIALDLNPFV 180
Qy 181 AKIPSPAISSQEQEPLLSVYAQAANLHLLLRDASIFGAEWGFTPGBEISTFYDQVTRTAQ 240
Db 181 SSIPSPAVSGHEVLLAVYAQAANLHLLLRDASIFGEEMGFTPGBEISFYNRQVLTAE 240
Qy 241 YSDYCVKNTWGLDKLKGNTAASWLYKHQFREMILLVLDVALPNYDTRYPYETTTAQ 300
Db 241 YSDYCVKNTWGLDKLKGNTAASWLYKHQFREMILLVLDVALPNYDTRYPYETTTAQ 300
Qy 301 LTRREVYTDPIVFNRETSGGFCRRWSLNSDISSEVESAVIRSPHFDILSEIEFFVTRAG 360
Db 301 LTRDVYTDPIAVNIVTSGFCRWSLNSDISSEVESAVIRSPHFDILSEIEFFVTRAG 360
Qy 361 LPLNNTEYLEYVWGHISIKYNT-NASSALERNYGTITSNKIKYDLANKDIFQVRSGLAD 419
Db 361 ITLNDAYINWNGHTLKYRRRTADSTVTANYGRITSEK-NSFALEDREDIFEINSTVAN 419
Qy 420 LANYAQAUVGPVYASTFLDKNTGSGVGFYTSKPHTTMOVCTQNYNTIDIPPENE-P 478
Db 420 LANYAQAUVGPVYASTFLDKNTGSGVGFYTSKPHTTMOVCTQNYNTIDIPPENE-P 478
Qy 479 LRGVSHRLSHITSYSGFNAGSPARYGNLPVPAWTHRSADVNTVYSDKITQIPWKAH 538
Db 478 VAEYSYHRLSHITSYSGFNAGSPARYGNLPVPAWTHRSADVNTVYSDKITQIPWKAH 538
Qy 539 TLVSGTIVKPGFTGNTLKTSSGFLAYTSVSKSPLSQRYRARIYASTTNLRLFTV 598
Db 536 MLVYGGVQVQPGFTGNTLKTSSGFLAYTSVSKSPLSQRYRARIYASTTNLRLFTV 598
Qy 599 ISGTRIYSINVNKTNMKGDDLPNTFDLATIGTAFPSNYSLTVGADSPASGGEVYVD 658
Db 595 YLGDTTEKRNKNTMDNGASLTETFKFASFTDFQRETQDKILLSMGDFSSGQEVYD 654
Qy 659 KPELIPVNATFEAEEDLDVAKAVNGLFTSKDALQTSVTDYQVNOAANLVECLSDLYP 718
Db 655 RIEFIPVDETEAEQDLAKAVNALFTNKDGLPGVTDYEVNOAANLVECLSDLYP 714
Qy 719 NEKRLMWDVAKEAKRLVQARNLQDTGPNRINGENGWGTGSTEIEVAGDVLFKDRSLRT 778
Db 715 NEKRLMWDVAKEAKRLVQARNLQDTGPNRINGENGWGTGSTEIEVAGDVLFKDRSLRT 778
Qy 779 SAREIDETPTLYQOIBESLLKPYTRYKLGFGIGSSQDLLEIKLIRHRANQIVKQVDPN 838
Db 775 GAREIDETPTLYQOIBESLLKPYTRYKLGFGIGSSQDLLEIKLIRHRANQIVKQVDPN 834
Qy 839 LLPDLVLPVNSCGGIDRCSEQYVDANLALENNENGENGMSDHSAPFHDITGEIDLNENT 898
Db 835 LLPDLVLPVNSCGGIDRCSEQYVDANLALENNENGENGMSDHSAPFHDITGEIDLNENT 888
Qy 899 GIWVVPKIPPTNGYATLGNLEVEEGLSGETLERAQOQEQQWQDQVARKRGASEKAYYA 958
Db 889 GIWVVPKIPPTNGYATLGNLEVEEGLSGETLERAQOQEQQWQDQVARKRGASEKAYYA 948
Qy 959 AKQAIIRLADYQDQKLVNSGVMSDMLAQNLVQSTPYVYNDALPIPGWNTYSPFELTN 1018
Db 949 SKQAVRLYADYQDQKLVNSGVMSDMLAQNLVQSTPYVYNDALPIPGWNTYSPFELTN 1008
Qy 1019 RLQQAANLVDLRNAIPNGDFRNLSDWNATSDVNVQOLSPTSVLVLPNWNVSQOQFTVQ 1078
Db 1009 RLQQAANLVDLRNAIPNGDFRNLSDWNATSDVNVQOLSPTSVLVLPNWNVSQOQFTVQ 1068
Qy 1079 PNRYVYLRVTRARKEGVDGYVIRIDGANTQETLTFNICDDDTGVLSADQTSYITKTVEFT 1138
Db 1069 PNRYVYLRVTRARKEGVDGYVIRIDGANTQETLTFNICDDDTGVLSADQTSYITKTVEFT 1128
Qy 1139 PSTEQWIDMSMETEGVFNIESVELVLEE 1167
Db 1129 PYTDQWIBISETEGTFIESVELVLEE 1157

```

RESULT 5

```

US-07-915-203-2
; Sequence 2, Application US/07915203
; Patent No. 5359048
; GENERAL INFORMATION:
; APPLICANT: Ohba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Reiichi
; APPLICANT: Suzuki, No. 5359048ukazu
; APPLICANT: Ogiwara, Katsutoshi
; APPLICANT: Sakanaka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawauchi, Tadaaki
; TITLE OF INVENTION: No. 5359048el Microorganism and Insecticide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/915,203
; FILING DATE: 19920723
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/K 301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-915-203-2

```

```

Query Match 60.6%; Score 3662.5; DB 1; Length 1149;
Best Local Similarity 63.8%; Pred. No. 4.6e-260;
Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

Qy 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDOTTLLQNMNYKDYLRKSEGENPELPGNP 60
Db 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDOTTLLQNMNYKDYLRKSEGENPELPGNP 60
Qy 61 ETFISS-STVQIGIGVGOVLGALGVPPAGQIASFVSFVQGLWPSSTSVVNMIMKQVE 119
Db 61 GTFISAQDAVGIGIDIVSTIISGLGIPVLGEVFSILGSLGLMPSNNENWQIFNVRVE 120
Qy 120 DLIDQKITDSVRKTALAGLQGLDGLDVYQKSLKNWLENRNDTRARSVVVTVIALELDF 179
Db 121 ELIDQKITDSVRKTALAGLQGLDGLDVYQKSLKNWLENRNDTRARSVVVTVIALELDF 180
Qy 180 VAKIIPSPAISSQEQEPLLSVYAQAANLHLLLRDASIFGAEWGFTPGBEISTFYDQVTRTA 239
Db 181 RTNMGFSQNTYETPLLPYQAASLHLLVMDVQIYKGEWGPQNDIDL FYKEQVSYTA 240
Qy 240 QYSDYCVKNTWGLDKLKGNTAASWLYKHQFREMILLVLDVALPNYDTRYPYETTTAQ 299
Db 241 RYSDYCVKNTWGLDKLKGNTAASWLYKHQFREMILLVLDVALPNYDTRYPYETTTAQ 300
Qy 300 QLTREVYTDPIVFNRETSGGFCRRWSLNSDISSEVESAVIRSPHFDILSEIEFFVTRAG 349
Db 301 EUTREIFTDV-----GSYVTGQSSTLISWYDMIPALPSPFTLEN-LLRKPDPFTLL 352

```

350 SEIEFYTRAGLPLNNT-EYLEYVWGHISIKYKNTNASSALERNYGTITSNKIKYIDLANK 408
353 QEIRMYTS---FRQNGTIEYNTWGGQRLTSLYIGSSP--NKYSGLVAGAEIDIIPVQON 407
409 DIFQVRLSGDLADLANYYAQQVGPYASFTLLDKTGTSGSGGFTYSPKPHHTMQVCTQYNT 468
408 DIYRV--VMTYIGRYTNSLGLVNPVTF-YFSNNTQK-----TYSKP-----KOFAGGIKT 454
469 ID---EIPPENEPLSRGSHRLGHTSYSPSKNASSPARYGNLPFAWTHRSADVNTVY 525
455 IDSGEELTYEN---YQSYSHRVSYITSPEIKSTGCTV--LCVPIFGWTHSSASRNFIY 509
526 SKITQIPVVAHTLVSGTIVIKPGP--FTGNIKIKRTSSQPLAY---TSVSVKSPISQRY 591
510 ATKISQIPINKASRTSGAVNFOEGLYNGGPPMKLSGSGSQVINLRVATDAKG-ASQRY 568
582 RARIVASTNLRFLVTIS-----GTRIYSINVNKWNKGGDLTENTEDLATIGTA-F 633
569 RIRIVASDRAGKF--TISRSRPNPATYSASIAYTNTMTNASTLYSTFAYAESGPINL 626
634 TFSNYSDSLTVGADSPAGSEVYVDPKELIPVNATFEAEEDLDVAKKAVNGLTFSKKDAL 693
627 GISGSRTPDISITKEAGANLYIDRIEFPVNTLFEAEEDLDVAKKAVNGLTFSKKDAL 686
694 QTSVTDVQVNOAANLVECLSELYPNKRMMLDAVKEAKLVQARNLLQDTGPNRINGEN 753
687 QTSVTDVQVNOAANLVECLSELYPNKRMMLDAVKEAKLVQARNLLQDTGPNRINGEN 746
754 GWTGSGIEVAGDVLFPKSLRLTSAREIDTETPTLYLYQOIDESEILKPYTKLKGFI 813
747 GWTGSGIEVAGDVLFPKSLRLTSAREIDTETPTLYLYQOIDESEILKPYTKLKGFI 806
814 GSSQDLKILIRHRANQIVKNVNDNLPLDVLVNSCGGIDRCSEQQVVDANLALNNGEN 873
807 GSSQDLKILIRHRANQIVKNVNDNLPLDVLVNSCGGIDRCSEQQVVDANLALNNGEN 866
874 GNMSSDSHAFSHIDTGEIDLNTNGIWWVFKIPTTNGYATLGNLELVEGSPISGETLER 933
867 GNMSSDSHAFSHIDTGEIDLNTNGIWWVFKIPTTNGYATLGNLELVEGSPISGETLER 926
934 AQOQEQQDMARKKASAKAYAAQADRLFPADYQDOKLNSGVMSDMLAQNILVQS 993
927 AQOQEQQDMARKKASAKAYAAQADRLFPADYQDOKLNSGVMSDMLAQNILVQS 986
994 IPYVNDALPEIPGMNYTSTFELTNRLQAWNLVQLNAPNGDFRNLGSDMNATSDVNV 1053
987 IPYVNDALPEIPGMNYTSTFELTNRLQAWNLVQLNAPNGDFRNLGSDMNATSDVNV 1046
1054 QOLSDTSLVILPKNNSQVSOQFTVQPNRYVLRVTKARKEGVDGYVIIRGANQTETLTF 1113
1047 QOLSDTSLVILPKNNSQVSOQFTVQPNRYVLRVTKARKEGVDGYVIIRGANQTETLTF 1106
1114 NICDDDTGVLSDQTSYITKTVEPTPTEQVWIDMSETE 1152
1107 NICDDDTGVLSDQTSYITKTVEPTPTEQVWIDMSETE 1145

RESULT 6

US-08-272-887-2
; Sequence 2, Application US/08272887
; Patent No. 5747450
; GENERAL INFORMATION:
; APPLICANT: Ohba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Reichi
; APPLICANT: Suzuki, No. 5747450kazu
; APPLICANT: Ogiwara, Katsutoshi
; APPLICANT: Sakanaka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawasaki, Tadaaki
; TITLE OF INVENTION: No. 5747450el Microorganism and Insecticide

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSES: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/272,887
FILING DATE: 08-JUL-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/915,203
FILING DATE: 23-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: M/K 301
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-272-887-2
Query Match 60.6%; Score 3662.5; DB 1; Length 1149;
Best Local Similarity 63.8%; Pred. No. 4.6e-260;
Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;
QY 1 MSPNNQNEYIILDASSSTVSNSRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60
DB 1 MSPNNQNEYIILDASSSTVSNSRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60
QY 61 ETPISS-STVQTGIGIVGVLGALGVPPFAGQATASFYSFVIGQLWPSSTVSWMINKQVE 119
DB 61 GTFISAQAVGTGIDIVSTIISGLGIPVLGEVFSILGSLGLLWPSNNENVMQIFMNRVE 120
QY 120 DLIDQKITDSVRKTALAGLQGLGDVYQKSLKNLENRNDTRASVVVTVQYIALELDF 179
DB 121 ELIDQKILDSVRSRAIADLANSRIVAYEYQNALEDWRKPHSTRSAALVKERFGNAEAIL 180
QY 180 VAKIPSPAISGOEVPVLLSVYAQANLHLLLRDASIFGAEWGTTPGEISTFYDRQVTRTA 239
DB 181 RTNMGFSQNTYETPLPTTYAQAASLHLLVMDVQIYKGEWGPQNDIDL FYKEQVSYTA 240
QY 240 QYSDYCVKXWYGLDKLKTNAASLWKYHQFREMTLLVLDLVALPNYDTRYPYIETTA 299
DB 241 RYSDHCQWYNAGLNKLRGTGAQWVDYKRFREMMVMVLDLVALPNYDTRYPYIETNA 300
QY 300 QLTREVTYDPIVENRETSGFCRRWSLN-----SDI-----SPSEVESAVIRSPHLPDIL 349
DB 301 ELTREIFTDPV-----GSYVTGQSSTLISWYDMIPAAALPSPSTLEN-LLRKDPPTLL 352
QY 350 SEIEFYTRAGLPLNNT-EYLEYVWGHISIKYKNTNASSALERNYGTITSNKIKYIDLANK 408
DB 353 QEIRMYTS---FRQNGTIEYNTWGGQRLTSLYIGSSP--NKYSGLVAGAEIDIIPVQON 407
QY 409 DIFQVRLSGDLADLANYYAQQVGPYASFTLLDKTGTSGSGGFTYSPKPHHTMQVCTQYNT 468
DB 408 DIYRV--VMTYIGRYTNSLGLVNPVTF-YFSNNTQK-----TYSKP-----KOFAGGIKT 454
QY 469 ID---EIPPENEPLSRGSHRLGHTSYSPSKNASSPARYGNLPFAWTHRSADVNTVY 525

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:10:20 ; Search time 30 Seconds
(without alignments)
3741.850 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYIILDASSSTSV.....MSETGCVFNIESVHLEEE 1167
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3733.5	61.8	1160	2 I40589	parasporal crystal
2	3123	51.7	1157	1 S49247	parasporal crystal
3	2750.5	45.5	1154	2 S39536	parasporal crystal
4	2547.5	42.1	1228	2 S00873	parasporal crystal
5	2458.5	40.7	1138	2 A48944	parasporal crystal
6	2191	36.3	1156	2 S19306	parasporal crystal
7	2154.5	35.6	1174	2 A42459	parasporal crystal
8	2093.5	34.6	1166	2 S32645	parasporal crystal
9	2064.5	34.2	1155	2 A26513	parasporal crystal
10	2061.5	34.1	1155	2 JD0002	parasporal crystal
11	2052.5	34.0	1155	2 S02134	parasporal crystal
12	2048	33.9	1176	2 J0241	parasporal crystal
13	2045	33.8	1176	2 JC2219	parasporal crystal
14	2044.5	33.8	1181	2 A41052	parasporal crystal
15	2043	33.8	1156	2 A29125	parasporal crystal
16	2038	33.7	1176	2 S02215	parasporal crystal
17	2037	33.7	1176	2 A22617	parasporal crystal
18	2030.5	33.6	1155	2 I39838	parasporal crystal
19	2024.5	33.5	1171	2 I40572	parasporal crystal
20	2010.5	33.3	1171	2 S37829	parasporal crystal
21	1989.5	32.9	1174	2 A32649	parasporal crystal
22	1984	32.8	1189	2 S00944	parasporal crystal
23	1953.5	32.3	1160	2 S32647	parasporal crystal
24	1944	32.2	1165	2 S11446	parasporal crystal
25	1939	32.1	1178	1 USBXKH	parasporal crystal
26	1932	32.0	1156	2 A29838	parasporal crystal
27	1930.5	31.9	1177	2 A49785	parasporal crystal
28	1923.5	31.8	1176	2 A48970	parasporal crystal
29	1917	31.7	1172	2 S32689	parasporal crystal

30 1868.5 30.9 1136 1 USB881 parasporal crystal
31 1778.5 29.4 1180 2 I39870 parasporal crystal
32 1772.5 29.3 1180 2 A26858 parasporal crystal
33 1734.5 28.7 719 2 I40590 cryV465 protein - insecticidal prote
34 1722.5 28.5 719 2 I39815 parasporal crystal
35 1713.5 28.4 719 2 S25383 insecticidal prote
36 1707.5 28.3 719 2 I39814 parasporal crystal
37 1683.5 27.9 934 2 A22798 parasporal crystal
38 1413 23.4 823 2 S04181 parasporal crystal
39 1372.5 21.1 652 2 I39811 parasporal crystal
40 1258.5 20.8 659 2 S10228 parasporal crystal
41 1245 20.6 652 2 A27323 parasporal crystal
42 1132.5 18.7 934 2 B29838 parasporal crystal
43 1116.5 18.5 649 1 JH0261 parasporal crystal
44 1069.5 17.7 655 2 JC7140 protoxin - Bacillu
45 1061 17.6 380 2 B42459 hypothetical prote

ALIGNMENTS

RESULT 1
I40589
parasporal crystal protein cry8Cal - Bacillus thuringiensis
N;Alternate names: parasporal crystal protein cryIII
C;Species: Bacillus thuringiensis
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 01-Dec-2000
C;Accession: I40589
R;Sato, R.; Takeuchi, K.; Ogiwara, K.; Minami, M.; Kaji, Y.; Suzuki, N.; Hori, H.; Asan
Curr. Microbiol. 28, 15-19, 1994
A;Title: Cloning, heterologous expression, and localization of a novel crystal protein
A;Reference number: I40589; MUID:94100786; PMID:7764305
A;Accession: I40589
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1160 <RES>
A;Cross-references: EMBL:U04366; NID:G532523; PIDN:AAA21119.1; PID:G532524
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 61.8%; Score 3733.5; DB 2; Length 1160;
Best Local Similarity 64.2%; Pred. No. 1.6e+208;
Matches 767; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

QY 1 MSPNNQNEYIILDASSSTSVSDNSRYPLANDQTTLQNNYKYDYLMSGENPELFCNP 60
DB 1 MSPNNQNEYIILDASSSTSVSDNSRYPLANDQTTLQNNYKYDYLMSGENPELFCNP 60
QY 61 EYFIS-STVQIGIGVQVGLGVPFAGQIASFYSPVIGOLWPSSTVSWEIMKQVE 119
DB 61 GTFISAQDAVGTGIDIVSTIISGLGIPVLGEVFSILGLLWPSNNVWQIFMNRVE 120
QY 120 DLIDOKITSVRKATAGLQGLDGVYQKSLKNWLENNDTRARSVVVQYIALELDF 179
DB 121 ELIDQKILDSVRSRATLANSRIAYEYQNALDWRKNPHSTRSAALVKERFGNAAIL 180
QY 180 VAKIPSAISGQVPLLSVYAQANLHLRLDASIFGAEWGTPGCHISTFYDQVTRTA 239
DB 181 RTWMSFSQNTYETPLPTVYAQAASHLLAVMRDVQVIGKEWGPONDILDPYKEQSVTA 240
QY 240 QVSDYCVKMYNTGLDKLKGTAASWKYHQFREMTELLVLDLVALFNYDTRTPYETTA 299
DB 241 RYSDHCQVQWYAGLNGTGRQWYDYNRFREMNVMVLDLVALFNYDARIYPLETNA 300
QY 300 QLTREVVYTPVFNRETSGFCRRMSLN-----SDI-----SFSEVESAVIRSPHLPDIL 349
DB 301 ELTREIFTPDV-----GSIVTQSQSTLISWYDMIPALPSPSTLEN-LLRKPDPFTLL 352
QY 350 SEIEFTYTRAGLPANN- EYLEYVWGHISIKYKNTASSALERNYGTITSNKIKYDYLANK 408
DB 353 QELRWYTS---FRQNGTIEYVWGGORLTLVYIGSSF--NKYSGVLAGAEDIIPVQGN 407
QY 409 DIPQVRSGLADLANYYAQVGVYPYAFSTLLDKMTGSGVSGGFTYYSKPHITNQVCTQNT 468

```
Db 408 DIYRV--VMTYIGRYTNSLLGVNPTP--YFSNNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPENEPISRGYSHRLSHLTSYSPKNSASPARYGNLVPFAMTHRSADVNTVY 525
Db 455 IDSGBELTYEN---YQSISHRVSYITSPEIKSTGTV--LGVPPIFGWTHSASRNNFIY 509
Qy 526 SKDIQIPIVVKARTLVSTGVTKGRG--FTGQNLKRTSGPLAY---TSVSVKSPLSORY 581
Db 510 ATKISQIDINKASRTSGCAVWNPFGELYNGGPMKLSGSGSQVIMLRVATDAKG--ASQRY 568
Qy 582 RARTRYSATTNLRFLFTTIS-----GTRIYSINVKNTMKRGDDLTFTNFDLATIGTA-F 633
Db 569 RIRIRYASDRACKP--TTSRSRPNPATYSASIAVTNTMSTNASUTYSTFEAVBSGPINL 626
Qy 634 TFSNYSDSLTVCADSFASGGEVYVDKFEIIPVNATFEABEEDLDVAKKAVNGLFTSKDIAL 693
Db 627 GISGSSRFPDISITKEAGANLYIDRIEIPVNTLFEABEEDLDVAKKAVNGLFTKEDAL 686
Qy 694 QTSVTDYQVNOQANLVECLSDBELYPNEKRMMDAVKAGRLVQARNLLQDTGTFNRINGEN 753
Db 687 QTSVTDYQVNOQANLIECLSDBELYPNEKRMMDAVKAGRLVQARNLLQDTGTFNRINGEN 746
Qy 754 GWTGSGTGLEVARGDVLKDRSLRLTSAREIDTETPTLYLQOI DESLLKPYTRYKLKGF 813
Db 747 GWTGSGTGLEVARGDVLKDRSLRLTSAREIDTETPTLYLQOI DESLLKPYTRYKLKGF 806
Qy 814 GSSQDLKELRHRANQIVKQNPVNDLLPDVLPVNSCGGIDRCSBQQYVDANLALNNGEN 873
Db 807 GSSQDLKELRHRANQIVKQNPVNDLLPDVLPVNSCGGIDRCSBQQYVDANLALNNGEN 866
Qy 874 GNMSSDSHAFSPHIDTGEIDANENTGIWVFKIPITNGVATIGNLELVEBGLSGTLEW 933
Db 867 GNMSSDSHAFSPHIDTGEIDANENTGIWVFKIPITNGVATIGNLELVEBGLSGTLEW 926
Qy 934 AQQQEQOQODKWARKEGSEKAYAAKQAI DRLPADYQDOKLNSGVMSDMLAQNLYQS 993
Db 927 AQQQEQOQODKWARKEGSEKAYAAKQAI DRLPADYQDOKLNSGVMSDMLAQNLYQS 986
Qy 994 IPYVYNDALPEIPGKNYTSFTELTNRLQOANLYDLRINAIPNGDFRNLSDNNATSDVNV 1053
Db 987 IPYVYNDALPEIPGKNYTSFTELTNRLQOANLYDLRINAIPNGDFRNLSDNNATSDVNV 1046
Qy 1054 QQLSDTSVLVIPNNNSQVSQQTQPNRYVLRVTRARKEGVGDGVVIRIDGANQTETLTF 1113
Db 1047 QQLSDTSVLVIPNNNSQVSQQTQPNRYVLRVTRARKEGVGDGVVIRIDGANQTETLTF 1106
Qy 1114 NICDDDTGVLSDADQSYITKTVEFTPSTEQWIDMSETGVNIESVELVLEE 1167
Db 1107 NICDDDTGVLSDQTSYITKTVEFTPSTEQWIDMSETGVNIESVELVLEE 1160

RESULT 2
S49247
parasporal crystal protein cry9Cal [validated] - Bacillus thuringiensis
N:Alternate names: parasporal crystal protein cryIH
C:Species: Bacillus thuringiensis
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C:Accession: A59350; S49247
R:Lambert, B.; Buysse, L.; Decock, C.; Janssens, S.; Pien, C.; Saey, B.; Seurinck, J.; V
Appl. Environ. Microbiol. 62, 80-86, 1996
A:Title: A Bacillus thuringiensis insecticidal crystal protein with a high activity agai
A:Reference number: A59350; MUID:96141404; PMID:8572715
A:Accession: A59350
A:Molecule type: DNA
A:Residues: 1-1157 <LAW>
A:Cross-references: EMBL:237527; NID:9547554; PIDN:CAA85764.1; PID:9547556
A:Experimental source: serovar tolworthi
C:Comment: This parasporal crystal protein, active against corn borer and other insects,
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 51.7%; Score 3123; DB 1; Length 1157;
```

```
Best Local Similarity 52.3%; Pred. No. 4,4e-173;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;
Qy 1 MSPNQNEYEILDASSSTSVSDNSVRYPLANDQTTTLQNMNYKDYLRMSEGENPELFGNP 60
Db 1 MNRNNQNEYEILDAPHCPCPSDDDDVRYPLASDPNAALQNMNYKDYLOMTDEDYDVSINP 60
Qy 61 ETPFSS-STVQIGIUGQVIGALGVPPAGQIATSPISFIVGQILMSPSSSTSVVEMIMKQVE 119
Db 61 SILSISGRDAVQTALTVVGRILGALGVPPSGQIVSFYQFLNLTLPVNDTAEIWEAFNRQVE 120
Qy 120 DLIDQKITDSVRKTAGLQIGLGDGLVYQKSLKWLNRNDTRARSVVVTVVIALELDF 179
Db 121 ELVNQOITTEFARNQALRLQGLGDSFNVTQRLQNLADRNDTRNLSSVVRAGFIALDLDF 180
Qy 180 VAKIPSAISQGEVPLLSVYQAANLHLLLDASIFGAEMGTFPGEISFTFYDQVTRTA 239
Db 181 VNAIFLFAVNGQVPLLSVYQAANLHLLLDASIFGEGMGTFQGEISFTFYDQVTRTA 240
Qy 240 QYSDYCVKMYNTGLDKLKGNTAASMLKYHOFREMTLLVLDLVALPPNTYTRTYPIETTA 299
Db 241 KYTNYCETMYNTGLDRLGNTNTESWLRYHOFREMTLLVLDVVALFPYDYVRLYPTGSPN 300
Qy 300 QLTREVYTDPIVFNRETSGGFCRRKSLNSDISFSVESAVIRSPHLFDILSEIEFYTTTRA 359
Db 301 QLTREVYTDPIVFNPPANVGLCRMGNTPYNTFSELENAFIRPPHLPDLRLNSLTISNR - 359
Qy 360 GLPLANTLEYLVWGHISIKYKNTNASSALLERNYGTITSNKIY---YDLANKDIFQVRS 416
Db 360 -FPV-SSNFMVMSGHTLRRSYLNSDAVQEDSYGLITTTTATINPGVDGTR-----IBST 413
Qy 417 GADLANYTAVYGVPPYASFTL--LDKNTGSGVGGFTYSKPHYTMQVCTQNTYTIIDIPP 474
Db 414 AVDFRSALIGYVGNRASFPVPGGLFNGTTPSPANG-----CRDLYDTNDELPP 461
Qy 475 ENEPISRGVS-HRLSHITSYSPSKN-ASSPARYGNLVPFAMTHRSADVNTVYSDKITOI 532
Db 462 DE---STGSSHLRSHVTFPFSQTNQAGSIANAGSVPTVWTRERDVLDANTITPNRITQL 518
Qy 533 PVVKARTLVSGTIVKGPFGTCGNILKRTSSGPLAYTSVSVKSPLSQRYARIRYASTTN 592
Db 519 PLVKASAPVSGTIVKGPFGTCGNILKRTSSGPLAYTSVSVKSPLSQRYARIRYASTTN 578
Qy 593 LELPVTISGTRIYSINVKNTMKRGDDLTFTNTP---DLATIG---TAPFTSNYSDSITVGA 646
Db 579 FSRILRGVSGIDVRLGSTMNRGQELTYESFTREFTTGTGPFNPPFTTQAQELITVNA 638
Qy 647 DSPASGGEVYVDKFEIIPVNATFEABEEDLDVAKKAVNGLFTSKDIALQTSVTDYQVNOQA 706
Db 639 EGVSTGGEYIIDRIEIPVNPAREABEEDLEAKKAVASLFTTRDGLQVNVTDYQVDOQA 698
Qy 707 NLVECLSDBELYPNEKRMMDAVKAGRLVQARNLLQDTGTFNRING--ENGWGTSGTGIEVA 764
Db 699 NLVSCLSDBEQYGHDKKMLLEAVRAAKRLSRERLLQDPDPTNTINSTEENGWKSNGVTIS 758
Qy 765 EGDVLPKDRSLTSAREIDTETPTLYLQOI DESLLKPYTRYKLKGFSGSODLRIKLI 824
Db 759 EGGPFKGRALQASAR---ENYPTIYQKVDASVLKPEYTRYLRLDGFVASSQDLIDLI 814
Qy 825 RHRANQIVKQNPVNDLLPDVLPVNSCGGIDRCSBQQYVDANLALNNGENGMSSDSHAPS 884
Db 815 HHKHLVKNVNDLVSDTYSVSGSCGINRCDEQHQVMDQLDAEHPHMDCEAAQTHBPS 874
Qy 885 FHIDTGEIDNENTGIWVFKIPITNGVATIGNLELVEBGLSGTLEWRAQQEQSQWQDK 944
Db 875 SYINTGDLNASVDQGIWVVKVTRTDGVTATLGNLELVEBGLSGTLEWRAQQEQSQWQDK 934
Qy 945 MARKGASEKAYAAKQAI DRLPADYQDOKLNSGVMSDMLAQNLYQVSIPIVYNDALPE 1004
Db 935 LGRKAEIDRVYLAQAQAINHLFVDYQDQQLNPEITGLABINEASNLVBSISGVSTDLQ 994
Qy 1005 IPGMNYSFTELTNRLQOANLYDLRINAIPNGDFRNLSDNNATSDVNVQQLSDTSVLVI 1064
```

Db 995 IPGINVEIYITELSDRLQQAQSYLYTSRNVAVQGNDSGLSDSWNTTMDASVQDQGNMFLVL 1054
Qy 1065 PNWNSQVSQQTVPQNYRYLVRTARKEGVGDGVYIIRGANOETETLTFNICDDDTGVLS 1124
Db 1055 SHWDAQVSQQLRVNPNKQYLVRTARKVGGGQGVYIIRGANOETETLTFNACDYVNGTY 1114
Qy 1125 ADQTSYITKVBTPSTPEQWIDMSETEGVFNIESVELVLEEB 1167
Db 1115 VNDNSYITBVFYFETKHMWVEVSESGSYIDSIPIETQEB 1157

RESULT 3
S39536
Parasporal crystal protein cry9Ba1 - Bacillus thuringiensis
N;Alternate names: delta-endotoxin-related protein; parasporal crystal protein cryX
C;Species: Bacillus thuringiensis
C;Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 01-Dec-2000
C;Accession: S39536
R;Shevelev, A.B.; Svarinsky, M.A.; Karasin, A.I.; Kogan, Y.N.; Chestukhina, G.G.; Stepan
FEBS Lett. 336, 79-82, 1993
A;Title: Primary structure of the cryX, the novel delta-endotoxin-related gene from Bac
A;Reference number: S39536; MUID:94085596; PMID:8262221
A;Accession: S39536
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1154 <SHE>
A;Cross-references: EMBL:X75019
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 45.5%; Score 2750.5; DB 2; Length 1154;
Best Local Similarity 48.7%; Pred. No. 1.9e-151;
Matches 562; Conservative 197; Mismatches 341; Indels 53; Gaps 18;

Qy 41 NYKDYLMSEGENPELFCNPETFTSSVTGIGVQVILGALGVPAGQIASFYSPIVG 100
Db 29 SYKDYLMSEGGYDYSINPGNV--RTGLQTGIDIVAVVVGALGPGVGGILTFGLTFG 86
Qy 101 QLPWSTVSVWEMTKQVEDLIDOKITDSVKTALAGLQGLDGLDVYQKSLKWLNRN 160
Db 87 FLWPSNDQAVWEAFIEQMEELIEQISDVVTRTALDDLTGTQNTYNYLYALKEWERPN 146
Qy 161 DTRARSVVVTQYIALELDFVAKIPSPALSGQ-----EVLPSVVAQAANLHLLLRASI 215
Db 147 GVRA-NLVLRQFETLHALFVSSMPSFG-SGPGSORFQAQLLVYVQAANLHLLLRABK 204
Qy 216 FGASWGTGPGISFTFYDROV-TRTAQYSDYCVKWNVTGLDKLKTNAASMLKYHQFRREM 274
Db 205 YGARWGLRESQIGNLYNELQTRDTYTHCVNAYNGLAGLGTSAESMLKYHQFRREA 264
Qy 275 TLLVLDLVALPPNYDTRTYPIETTAQLTREYVTDPIVFNRETSGP-----CRRWLSNDSI 330
Db 265 TLMAMDIALPPYNTNRRYPVAVNPQLTREYVTDPLGVPSSESLPPELRLCLRWQETSAM 324
Qy 331 SFSEVESAVIRSPHLPDILSIEBFTTFRAGLPIANNTEYLYWVGHSI--KYNTNASSAL 388
Db 325 TFSNLENAIISPHLFDITNNLMITYGFSFVHLTN-QLIEGWIGHSVSSLLASGPTVL 383
Qy 389 ERNYGTITSNKIKYVLDLANKDIFQVRS--LGAULANYAQQVYGVYVPSFATLLDKNTGSG 445
Db 384 RRTGSGTTS-IWTFSEFNDRVYQNTSRHTGLGFQX-APLFGITRAQF----- 430
Qy 446 SVGGFTYSKPHTTMQVCTQNTYNTIDEIP--PENEPLSRGYSHRLSHITSYSF-----SK 497
Db 431 -YPGTYSVTQNALTCQYNYSIDELPSLDPNEPISRSYSHRLSHITSYLYRLVLTIDGI 489
Qy 498 NASSPARYGNLPVANTHRSADVNTVYSDKITQIPVYKANTLYSGITVIKPGFTGGNI 557
Db 490 NIYS-----GNLPTVYVWTRHVDLTNTITADRITQLPLKSPFEPAGTIVVRGPGFTGDI 545
Qy 558 LKRTSSGPLAYTSVSVKSPQSVRIRYRIRASTTNLRLFVTSIGTRIYSINWNTMKNKG 617
Db 546 LRRTGVTGFGIRVTRTAPLQRIKIRFRFPASTTNLPIGIRVGRQVNYFPFGRTNMRGD 605

Qy 618 DLTPNTFDLATIGTAFPSNYSDSLTVCAQSPASGEVYVVDKPELIPYNATPEAEEDLDV 677
Db 606 ELAYESPATREFTDFNFQPELISVFANAFSAQGVYFDRIEIIPNPAKEAKEDLEA 665
Qy 678 AKKAVNGLFTSKKDALQTSVTDYVQNOAANLVECLSDELYPNKEKMLMDVAKEAKRLVQA 737
Db 666 AKKAVASLFTTRDLQVNVKDYVQQAANLVSCLSDQYGVDKMLLEAVRAAKRLSRE 725
Qy 738 RNLLQDTGFNRING--ENGWGTSGTIEVARGDVLFKORSLSLTSAEIDTETYPYTLVQO 795
Db 726 RNLLQDDPDTNTINSTENGWKAANGVTISSEGGPFYKGRALQALASAR-----ENYPTIYQK 781
Qy 796 IDESLAKPYTRYKLKPGTSSODLEIKLIRHRANQIVKGVDPDNLPLVLPVNSCGSIDRC 855
Db 782 VDASELPYTRYKSDGPFVKSQDLEILDIHHKRVHLVKVNPVNDLVSDTYPDDSCSGINRC 841
Qy 856 SEQQYVDANLALENKGE--NGMNSDSHAFSPHIDTGEIDLNENTGIWVVPKIPYTNQYAT 914
Db 842 QEQQWVAQLETEHHHPMDCEAAQTHFESSYIDTGLNLSVQDQIGIMAFKVRTTQGYAT 901
Qy 915 LGNLELVEGSLSGETTLERAAQQOQMODQWARKGASEKAYAAQAIDRLFPADYQOQK 974
Db 902 LGNLELVEGSLSGESLREQRDNTKWSAELGRKRAETDRVYQDAKQKQINHLFVDYQDQO 961
Qy 975 LNSGVMSDMLAAQNLVQSIPIYVYNDALPEIPGNYTTSFTLTNQLQANMLYDLRNAIP 1034
Db 962 LNPEIGHMADIMDAQNLVASISDVYSDAVLQIPGINYEIYTELSNRLQQAQSYLYTSRNVQ 1021
Qy 1035 NGDPFRNGLSDWNTSDVNVQQLSDTSVLVIPNMSQVSQQFTVQPNRYVYLVARTAKGV 1094
Db 1022 NGDPFRNGLSDWNTAGASVQDQGNTHFLVLSHMDAQSQQFVQPNCKYVLAARTAKGV 1081
Qy 1095 GDGVYIIRGANOETETLTFNFCDDDTGVLSDAQTSYITKTVEFTSTEQVWIDMSETEGV 1154
Db 1082 GDGVYIIRGANOETETLTFNACDVIINGTYVTDNTYLTKEVIFYSHTHEMVEVNETEGA 1141
Qy 1155 FNIESVELVLEEB 1167
Db 1142 PHIDSIEFPVETEK 1154

RESULT 4
S00873
parasporal crystal protein cryBa1 - Bacillus thuringiensis subsp. thuringiensis
N;Alternate names: parasporal crystal protein cryA4
C;Species: Bacillus thuringiensis subsp. thuringiensis
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 01-Dec-2000
C;Accession: S00873
R;Brizzard, B.L.; Whiteley, H.R.
Nucleic Acids Res. 16, 2723-2724, 1988
A;Title: Nucleotide sequence of an additional crystal protein gene cloned from Bacillus
A;Reference number: S00873; MUID:88203216; PMID:3362680
A;Accession: S00873
A;Molecule type: DNA
A;Residues: 1-1228 <BRI>
A;Cross-references: EMBL:X06711; NID:940264; PIDN:CAA29898.1; PID:9580949
C;Genetics:
A;Gene: cryA4
A;Start codon: TTG
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 42.1%; Score 2547.5; DB 2; Length 1228;
Best Local Similarity 43.8%; Pred. No. 1.3e-139;
Matches 558; Conservative 183; Mismatches 381; Indels 151; Gaps 21;

Qy 1 MSPNQNYEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLMSEGENPELPGNP 60
Db 1 MTSNRKNNEIINAHNSHA-----QMDLLPDARIEDSLCIAEG-----NNI 42
Qy 61 ETPFSSVTQGTGIGVQVILGALGVPAGQIASFYSFVIGQLWPSSTVSWEMIMKQVED 120

Db 43 DPFTSASTVQGINIAGRIILGVLPFAGQQLASFTSVFLUGELWPRGR-DQWEIEFLEHVEQ 101
Qy 121 LIDOKITDSVARTALAGLQGLDGLDYYQSKLKNWLENDRTRASVVVVVQYIALELDPV 180
Db 102 LINQOIITENARTALARLQGLGDSFRAQQSLEDWLENDRDARTSVLYTQYIALELDFL 161
Qy 181 AKIPSPAIQGEVPLLSVYQAANLHLLLDASIFGAENGFTGCEISTFYDQVTRTAQ 240
Db 162 NAMPLFAIRNQEVPLLMVYQAANLHLLLDASLFGSEFGLTQSEIQRYYERQVERTRD 221
Qy 241 YSDYCVKMYNTGLDKLKGNTAASWLKYHQFREMTELLVLDLVALFPNPDYDTRTYPIETTAQ 300
Db 222 YSDYCVKMYNTGLSLRGTNAASVWRVYQFRRDLTGLVLDLVALFPSPDYDTRTYPIETTAQ 281
Qy 301 LTRVYDTPDVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLDILSIBIFTYTTRAG 360
Db 282 LTRVYDTPDVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLDILSIBIFTYTTRAG 337
Qy 361 LPLNTEYLEVWGHSHIKYK-----WTNASSALERNVGTITSNKIKYDILANKDIFQVR 414
Db 338 SRWSENTRHMYRGTHTQSRPIGGLANTSTHGANTNTSINPVTLR-----FASRDVYRTE 391
Qy 415 SLGADL--ANYAAQVYGVYASFTLLDKNTGSGVGGFTYKSPHTTMQVCTQNTYNTIDEI 472
Db 392 SYAGVLLWGLYLEPIHGVTVRFTNPONISDR-GTANYSQPYESPGQLQKDSER--EL 448
Qy 473 PPE--NEPLSRGYSHRLSHITSYFSKNASSPARYGNLPVPAWTHRSADVNTVYSDKIT 530
Db 449 PPEETTERPNVYSYSHRLSHIGIILQSR-----VNVVYVSWTHRSADRTNTIGPNRIT 500
Qy 531 QIPVVKATLVSGTIVTKPGPTGCTGNTILKRTSSGFLATSVSVKSPLSQRVRAIRVAST 590
Db 501 QIPVVKASELPQGTIVVVRGPGTGGDILRRNTGCGPIRVTVNGPLTQRYGRFYAST 560
Qy 591 TNLRLFTVISTRIYSINVKNTMKGGDLDLNTDNLATIGTAFPTFSNYSLSLTGADSPA 650
Db 561 VDPDFVSRGTTVNNRFLRTMSGBELKYNFVRRAFTPTPTTQIQDIIRTSIOGLS 620
Qy 651 SSGEVYVDKPELIPVNAATFEABEDLVAKKAVNGLFTSKKD-ALQTSVTDYQVNNQANLV 709
Db 621 GNGEYWDIKBIIPVATFEAYDLERAQAVNALFTNPNRLKTDVTDYHIDQVSNLV 680
Qy 710 ECSDYELPYNEKMLDAVBAKLVQARNLLODTGNRIN----- 750
Db 681 ACUSDPEFLDSEKRELLEKVKYAKLSDERNLQDPNFTSINKQDPFITNEQSNFTSIHE 740
Qy 751 -GENGTGSGTIEVAGDVLFPKDRSLRSLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 809
Db 741 QSEHGHWGSENIITQEGNDVPEKENYVYLPQT---FNECYPYLYQKIGESSELKAYTRYQL 797
Qy 810 KGPTGSSODLEIKLIRHRANOIVRNP--DNLLPDVL--PVNSCGGIDRCSBQOYVDANL 865
Db 798 RGYTEDSQDLEIYLIRYNAKHETLDVPGTESLWPLSVESPTGRCCEPNRCA--PHFEWNP 855
Qy 866 ALENNENG-NMSSDSHAFSPHIDTGEIDLENNTGIVWVFKIPTNGVATIGNLELVEEG 924
Db 856 DLDCSDRGEKCAHSHHPSLDIDVGCTDLHENIGVWVFKIPTQEGHARLGNLEFTEEK 915
Qy 925 PLSGETLERAOOQOQWDMARXRGASAKAYAAKOAILDLFADYQDQKLSNVSMSDM 984
Db 916 PLLGEALSRVKAEEKWDRKLEKQLETKRYVYTEAKEAVDALFVDSQYDRLOADNTNIGMI 975
Qy 985 LAAQNLVQSIPIYVNDALPEIPGMNYSFTBLTNRLOOANWLYDLRNLAINPGDFRNLGSD 1044
Db 976 HAADKLVRHIREAYLSELFPVPGVNAEIFEBLEGHIIITALSIDARNVWKNDFPNGLTC 1035
Qy 1045 WNATSDVNVQOLSDTSVLVIPNWSQVSOQPTVQPNRYVLRVTRARKEGVDGVVIRDG 1104
Db 1036 WNVKGHVDVQOQSHRSDLVIFPEWBAEVSQAVRVCPGCCYILRLVYAYKEGVEGCVTIHEI 1095
Qy 1105 ANOTETLTF----- 1113
Db 1096 ENWTDLKFKNRBEESVYPTDTGTCDNTAHQGTAGCADACNSRAGYEDAYEYDVTASV 1155

Qy 1114 -----NICDDTGVLSAD--QTSYITKTVETPSTBQVMDIMSETBGV 1154
Db 1156 NYKPTYBEETVDRRNHNHCEYRGVYVNPVPVAGVYTKBLEYFPDETDTVMISIGETGCK 1215
Qy 1155 FNIESVBLVLEEE 1167
Db 1216 FIVDSVLELLMEE 1228
RESULT 5
A48944
parasporal crystal protein cry7Aa1 - Bacillus thuringiensis
N;Alternate names: parasporal crystal protein cryIIIC
C;Species: Bacillus thuringiensis
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C;Accession: A48944
R;Lambert, B.; Hofte, H.; Anys, K.; Jansens, S.; Soetaert, P.; Peferoen, M.
Appl. Environ. Microbiol. 59, 2536-2542, 1992
A;Title: Novel Bacillus thuringiensis insecticidal crystal protein with a silent activi
A;Reference number: A48944; MUID:92384571; PMID:1514800
A;Contents: BTS137J
A;Accession: A48944
A;Status: preliminary
A;Molecule type: DNA; protein
A;Residues: 1-1138 <LAM>
A;Cross-references: GB:M64478; NID:g142760; PIDN:AAA22351.1; PID:g142761
A;Note: sequence extracted from NCBI backbone (NCBIN:112092, NCBIPI:112093)
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 40.7%; Score 2458.5; DB 2; Length 1138;
Best Local Similarity 43.5%; Pred. No. 1.6e-134;
Matches 514; Conservative 208; Mismatches 400; Indels 59; Gaps 17;

Qy 1 MSPNQNEYELLDASSSTSVSDNSVRYPLANDQTTTLQNMVYKDYLRMSEGRNPELFGNP 60
Db 1 MNLNLDGYE-----DSNRTLNNSLNPTQALSPSKNNMYTQDPLSITERQPEALASG 55
Qy 61 EFTFISSTVQTGIGIVQVIGALGVPPAGQIATSYFIVGQLWFSSTVSVNEMIMKQVED 120
Db 56 NT-----AINTVSVTGATLSALGPGASPITNFYKLAGLLWPENG-KIWDEPMTVEA 109
Qy 121 LIDOKITDSVARTALAGLQGLDGLDYYQSKLKNWLENDRTRASVVVVVQYIALELDPV 180
Db 110 LIDOKIEYVRNTKAIABLDGLGSLDKYQKALADWLKGQDDPEALISVATFRIIDSLFE 169
Qy 181 AKIPSPAIQGEVPLLSVYQAANLHLLLDASIFGAENGFTGCEISTFYDQVTRTAQ 240
Db 170 FSWFSFKVTGYEIPLLTYQAANLHLLLDASLTDSTLYGDKGFTQNNIBENTYRQKR1SE 229
Qy 241 YSDYCVKMYNTGLDKLKGNTAASWLKYHQFREMTELLVLDLVALFPNPDYDTRTYPIETTAQ 300
Db 230 YSDHCTRWYNSGLSLNGSTVYBQWYNRFRREMILMALDLVAVPPPHDPHRRYSMETSTQ 289
Qy 301 LTRVYDTPDVFNRETSGGFCRRWSLNSDI--SFSEVESAVIRSPHLDILSIBIFTYTTR 358
Db 290 LTRVYDTPDVSLSSIS-----NPDIGSPFSQMENTAIRPHLDVLDLDELVIYTSK 338
Qy 359 AGLPLNTE-YLEVWGHSHIKYKNTNASSALERNVGTITSNKIK--YYDLANKDIFQVRS 415
Db 339 YKAPSHIQDPLFWSAHKVSFKKSESNLYTTGIYKTSYISGAYSFGNDIYRTLA 398
Qy 416 LGADLANYYAQVGVYPYASFTLLDKNTGSGVGGFTYKSPHTTMQVCTQNTYNTIDEIPEE 475
Db 399 APSVWVVPYQYQVGEQVEFY-----GVKGVHYRGDNKYDL---YDSDIDQLPDD 446
Qy 476 NEPLSRGYSHRLSHITSYFSKNASSPARYGN--LPVPAWTHRSADVNTVYSDKITIP 533
Db 447 GEPTHEKYTHRLCHATAI-----PKSTPDYDNATIPISWTHRSABEYRNPKNKITKP 501
Qy 534 VVKAHTLVSGTIVTKPGPTGCTGNTILKRTSSGFLATSVSVKSPLSQRVRAIRVASTNL 593
Db 501

Db 502 AVRMKLDLPSTVVKPGFTGGLVKGSTGYIGDIKATVNSPLSQRYVRVATNVSG 561
QY 594 RLFTVTSIG-----TRIVSVNNKTMKGGDLTFNTPLATIGTAFPTSNYSDSLTVGADS 648
Db 562 QFNVIYNDKITLQTKFO--NTVETIGEGKOLTYGPGYIYVSTTIQPPDBHPKTIHLSD 619
QY 649 FASGGEVYVDPFELIPVNAFTAEDEDLVAKAVNGLFTSKKDALQTSVTDYQVNOAANL 708
Db 620 LSNSSFTVDSIEPIPVVDVNAEKEKLEKAKQAVNTLFTGKNALQKVDYKVDQVSI 679
QY 709 VECISDELVEPNEKMLMDVAKKRLVQARNLQDTGFRNIRG--ENGWGTGSGTIEVAEG 766
Db 680 VDCISGLDLYPNEKELQNLVKYAKLSYSRNLLDPTFDSINSSENGWYGSNGIVGNG 739
QY 767 DVLPKORSRLTSAREIDTETYPYLYQQIDESLLKPYTRYKLGKFGTSSQDLLEIKLRH 826
Db 740 DFVFKGNLYIFSGTN--DTQ-YPTLYQKIDESKLKEYTRYKLGKFGTSSQDLLEAVYIR 796
QY 827 RANQIVKRVDPNLLPDLVPSNCGGIDRCSEQQVVDANLALNNGENGNWSSDSHSPFH 886
Db 797 DAKHRTLOVDNLLPDILPENTCGEPNRCAAQQYLDENPSPCCSMQDGLSDSHSFSLN 856
QY 887 IDTGEIDLNTGNTGWVFKPTPTNGYATLGNLELVERGPLSGETLERAAQQOEOQWQDMA 946
Db 857 IDTGSINHNENLGIWVLFKISTEGYAKFNLEVEDPGVIGEALARVKQETKWRKLA 916
QY 947 RKGASEKAYAAQADIRLPADYQDOQLNSGVMSDMLAAQNLVQSIPIVYNDALPEIP 1006
Db 917 QLATTETQAIYTRAKQALDNLPAQAQSHLKRDTVFABIAAAKIVQSIREAYNSWLSWP 976
QY 1007 GMYTSTELTNLQAAWNLVDLNAIPNGDFRNLGSDNATSDVNTVQQLSDTSVLVFPN 1066
Db 977 GVNHPIFTELSGRVQRAFQYDVRNVVRNGRFLNGLSDWITVSDVKVQOBENGNNVLVNN 1036
QY 1067 WNSQVSQOFTVQPNRYRVLTARKEGVGQYVIRDGANQOTETLTENICDDDTGVLSAD 1126
Db 1037 WDAQVLQNLVLYQDRGVLHVLTARKIGIGEGYITIDEEGHTQDLRTAC-EEIDASNAP 1095
QY 1127 QTSYITNTKVTFTSTEQVWIDMSTEGVFNIESVELVLEE 1167
Db 1096 ISGVITKELEFPFDPTEKVVHIEGTETGIFLVESIELEFLMEE 1136

RESULT 6
S19306
parasporal crystal protein cry9Aa1 - Bacillus thuringiensis
N:Alternate names: delta-endotoxin; insecticidal crystal protein; parasporal crystal pro
C:Species: Bacillus thuringiensis
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 01-Dec-2000
C:Accession: S19306; S23588; A44847; S14602; S14837
R:Smulevitch, S.V.; Osterman, A.L.; Shevelev, A.B.; Kaluger, S.V.; Karasin, A.I.; Kadyro
FEBS Lett. 293, 25-28, 1991
A:Title: Nucleotide sequence of a novel delta-endotoxin gene cryIlg of Bacillus thuringie
A:Reference number: S19306; MUID:52070568; PMID:1660003
A:Accession: S19306
A:Molecule type: DNA
A:Residues: 1-1156 <SMU>
A:Cross-references: EMBL:X58120; NID:g870929; PIDN:CAA41122.1; PID:g40271
A:Experimental source: subsp. galleriae
A:Accession: S23588
A:Molecule type: protein
A:Residues: 24-34 <SMU>
A:Experimental source: subsp. galleriae
R:Gleave, A.P.; Hedges, R.J.; Broadwell, A.H.
J. Gen. Microbiol. 138, 55-62, 1992
A:Title: Identification of an insecticidal crystal protein from Bacillus thuringiensis
A:Reference number: A44847; MUID:92211329; PMID:1556556
A:Accession: A44847
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1151 <GLE>
A:Cross-references: EMBL:X58534; NID:g48879; PIDN:CAA41425.1; PID:g48880
A:Experimental source: isolate DSIR517

A>Note: sequence extracted from NCBI backbone (NCBIN:92865, NCBIIP:92867)
C:Genetics:
A:Gene: cryIlg
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 36.3%; Score 2191; DB 2; Length 1156;
Best Local Similarity 41.5%; Pred. No. 5.7e-119; Indels 98; Gaps 24;
Matches 501; Conservative 195; Mismatches 414

QY 5 NQNEYELDASSSTVSVDNSVRYPLANDQTTTLLQNNYKQYLRMSGEENPBLFGNPETPI 64
Db 2 NQNKHGIIGASCCGASDDVAKYPLANNPYSSALNIN-----SC 40
QY 65 SSSTVQVGIGIVGQ-----VLGALGVP-FAGQIASFYFIVQLPSPSTVSWE 112
Db 41 QNSSILNWINIIGDAEKAESVIGTIVLSLTAPSLTGLISIVYDLIGKVLGSSQGISD 100
QY 113 MIMKQVEDLLDOKITDSVRKTALAGLQGLDGLDVYQKSLKNWLENDRDTRARSVVVTOY 172
Db 101 LSIICDLSIIIDLRVSQSVLNDGADFNQSVLLRYNYLEALDSMNKNPNASABERTRPR 160
QY 173 IA-LELDFVAKIPSAISGQ-----EVPILLSVYAOAANLHLLLRDASIFGABWGF--- 222
Db 161 IADSEFDRILTRGSLTNGGSLARQNAQIILLPSFASAAFPHLLLRDADTRVGNWGLYNA 220
QY 223 TPGEISTFYDQVTRTAQYSDYCVKWNTGLDKL--KGTNAAASMLKYHQPRREMTLLVLD 280
Db 221 TP--FINYQSKVELIELYTDYCVHWNRGFNELRQGTSTATAWLEFHYRREMTLWLD 278
QY 281 LVALPNNYDTRTYPIETTAQLTREVYDPIVFNRETSGGFCRRSLNSD-----TSF 332
Db 279 IVASPSLDTITNPIETDFQLSRVITYDPI-----GFVRSLSRGESWFSFVNRAFP 330
QY 333 SEVESAVIRSPHFLDILSEIEFTYTRAGLPLN-NTEYLEYVWVGHISIKYKTNASSALBRN 391
Db 331 SDLENA-IPNPRSPFLNMIISTGSLTLVPSPTDRARVWYGSRRDRISPANSPITELI 389
QY 392 YGITSNKIKYDLANKDIFQVRSGLADLANYYAQYGVVPVASTLDDKNTGS-----G 445
Db 390 SQGHIT---ATQITLGRNIIPVDSQACNLAD---TTYGVNRAVF-YHDASGSGRSVYEG 442
QY 446 SVGGFTYSKPHITMQVCTQNYNTIDEIPPENE--PLSRGYSHRLSHITSYSPSKNASSPA 503
Db 443 YIRTTGIDNPR-----VQINITY--LPGENSDIPTEDYTHILSTINLTGGLRQVASN 494
QY 504 RYGNLPVPAMTHRSADVNTVYSDKIQIPVKAHTLVSGTTVIKPGFTTCGNILKRTSS 563
Db 495 RRSLLWYGVTHKSLARNNTINPDRIQIPLTKVDTRGTVGSVYNDPGFIGGALLQRTDH 554
QY 564 GPLAYTSVSVKSPISQRYRARIYASTTNLRLFTVIS-GTRIYSINNVKNTNKGDDLTFFN 622
Db 555 GSLGVLRVQFPFLHRLQQRIRVRVASTTNIRLSVNGSFGT--ISQNLPTSMRLGSDLRYG 612
QY 623 TFDLATIGTAFPTSNYSDSLTVGADSPASGEVYVDKFEELIPVNATPAEEDLDVAKKAV 682
Db 613 SFAIREPNTSIRPTASPDQRLRTIEPSFIRQEVYVDRIEPIPVNPTREKSDLEAKKAV 672
QY 683 NGLFTSKKDALQTSVTDYQVNOAANLVECLSDLEYPNKRLMDVAIVEAKELVQARNLLQ 742
Db 673 ASLFTTRDGLQNVKQYQVDOAANLVCLSDEQGYDKKMLLEAVRAAKLSRENNLQ 732
QY 743 DTGFNIRING--ENGWGTGSGTIEVAAGDVLPKDRSLRSLTSAREIDTETYPYLYQQIDSSL 800
Db 733 DDPFNTINSTENGKWKASNGVTISSEGGPFYKGRALQASAR-----ENYPTIYQKVDASE 788
QY 801 LKPYTRYKLGKFGTSSQDLLEIKLRHRAQVIRKVPDNLPLDVLVNSCGGIDRCSEQY 860
Db 789 LKPYTRYKLGKFGTSSQDLLEIKLRHRAQVIRKVPDNLPLDVLVNSCGGIDRCSEQY 848
QY 861 VDNALALENNGE--NGNMSDSHAFSPHIDTGEIDLNTGNTGWVFKPTPTNGYATLGNLE 919
Db 849 VNAQLETEHHHPMDCCCAAQOTHEFFSSYIDTGLDNLSSVDQGIWAFKRVTTDGYATLGNLE 908

QY 1044 DNATSDVNV--QQLSDTSVLVIPNWSQVSOQFTVQPNYVVLVLTARKEGVGDGYVILR 1103
 DB 974 CMNVKGHVDVEBQNNHRSVLVVPWEAEARVSQEVRCFGRGYILRVITAYTEGVBEGCVTIH 1033
 QY 1103 DGANOTETLTFN-----ICDDDTG-----VLSA 1125
 DB 1034 EIENNTDELKFSNCVEEVVFNNTVTCNDYTAQOEYKGYATSHNRGDEAYGNPSPVA 1093
 QY 1126 DQT-----SYITKTVEPTPSTEQWIDMSETEGV 1154
 DB 1094 DYTVPVEKAYDGRRENPCSNRGYDYTPLPAGYVTKLEVFPETDKVWIEIGTEGT 1153
 QY 1155 PNIESVBLVLEE 1167
 DB 1154 FIVESVBLLEE 1166

RESULT 9
 A:parasporal crystal protein - Bacillus thuringiensis (strain aizawai)
 C:Species: Bacillus thuringiensis
 C:Date: 11-Mar-1988 #sequence_revision 11-Mar-1988 #text_change 01-Dec-2000
 C:Accession: A26513
 C:Oeda, K.; Oshie, K.; Shimizu, M.; Nakamura, K.; Yamamoto, H.; Nakayama, I.; O
 Gene 53, 113-119, 1987
 A:Title: Nucleotide sequence of the insecticidal protein gene of Bacillus thuri
 A:Reference number: A26513; MUID:97248103; PMID:3297927
 A:Accession: A26513
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1155 <GB>
 A:Cross-references: NID:g143098; PIDN:AAA22551.1; PID:g143099
 C:Superfamily: parasporal crystal protein
 C:Keywords: delta-endotoxin

Query Match 34.2%; Score 2064.5; DB 2; Length 1155;
 Best Local Similarity 39.0%; Pred. No. 1.3e-111;
 Matches 476; Conservative 184; Mismatches 387; Indels 173; Gaps 27;

QY 53 NPFL-----FCNPET-FISSSTVQTG---IGI---VGQVLGALGVPPAGQIASPSP 97
 DB 4 NPNINECIPNCLSNPNEVVGIERETGYTPTDLSLTFQLLSEVPGAGPVLGLVDI 63
 QY 98 IVGQLWPSSTVSVMETMKQVBOLIDQKITSVRKTALAGLQGLDGLVYKSLKNWLE 157
 DB 64 IWGIFGPSQ---WDAFLVQIEQLINQRIEFPARNOAISRLGSLNLYQIYAESPREWEA 119
 QY 158 NRNDTRASVVVTVQYIALELDLFWAKTSPASISQEVPLLSVYAQAANLHLLLRDASIFG 217
 DB 120 DPTNPALRBENRIQFNDMNSALTTAFLPFAVQNVPLLSVYVQAANLHLVSLRDSVFG 179
 QY 218 AEWGTFPGESTFYDQVTRTAQVSDYCVKWYNTGLDKLGTNAASLWKLKHQFRRMTLL 277
 DB 180 QRNGFDAAITNSRYNDLTRLIGMYTTHAVRWYNGLERVWGPDSRDWRINQFRELTLT 239
 QY 278 VLDLVALFPNYDRTYPIETTAQLTREVYTDPIVFNRETSGGFCRRWSLNSDISFSEVES 337
 DB 240 VLDIVSLFPNYSRTYPIRTVSQLTREIYNPVL-----ENFDGSRALAQ 285
 QY 338 AV---IRSPHLFDLSIEIFTTTRAGLPLNTHYLEYWUGHSTK-----YKNTNASSALE 389
 DB 286 GIEGIRSPHMLDILNSITTYT-----DAHRGEY--YWSGHQIMASVPVGFSPGFPTPL- 337
 QY 390 RNYGTITSNKIKYVDLANKDIQVRSIGADLANVYAQVY--GVPYASFTLLDKNT---G 443
 DB 338 --YGTMGNAAPQORIVAQLQGGVYRTLSSTL-----YRRPFNIGINNOQLSLVDLTGTFAYGT 392
 QY 444 SGVGGPTYSKPHTTMQVCTQNTYNTIDEIPPENE--PLSRGYSHRLSHITSY--SFSKNA 499
 DB 393 SSNLPSAVYKSGTV-----DSLDEIPQNNVPPRQCPSHRLSHVSMFSGFSNNS 444
 QY 500 SSPARYGNLPEVFAWTHRSADVTNTVYSDKTIQIPVVKAHTLVSGTTVIKQPGFTGNILK 559

Qy	503	ARYGNLPVFAWTHRSADVNTVTYSDKIQTQIPVWKAHTLVSGTWTWIKGPGFTGGMILKRTS	562
Db	448	IR---AMPFSIHRSABENNIIPSSQITQIPLTKSTNLGSGTVVVGPGFTGGDILRRTS	504
Qy	563	SGPLAYTSVSVKSPLSQRYARIIRYASTNLRLPVTISGTRIYSINVNKNKGGDDLTFN	622
Db	505	PGQISTLRNITAPLSQRYRIRIYASTNNLOFHTSIDGRPINOGNSFATSSGSSGQSG	564
Qy	623	TFDLATIGTAFTFNGYNSDSLTVGADSFASGCEVYVDKPELPVNAATFEAREEDLOVAKKAV	682
Db	565	SFRTVGTTTTFNFGSGSVFTLSHVFNSGNEVIDRIEFVPAEVTFEAEYDLERAQKAV	624
Qy	683	NGLFTSKDD-ALQTSVTDYQVNOAANLVBCLSDLEYLNEKRMKLWDVKEAKRLVOARNLL	744
Db	625	NELFTSSNOIGLKTVDYHIDQVSNLVBCLSDLEFCLDEKKEKSEKVHAKRLSDERNLL	684
Qy	742	QDTGPNRING--ENGWTGSGTCIEVABGDVLPKDRSLRLTSAREIDTETYPYTLVQOQDES	799
Db	685	QDNPRGINQLDRGWRGSTDITIQGGDDVFKENYVTLGLTFD---BCYPYLYQKIDES	741
Qy	800	LLKPYTRYKLKPGTSSQDLBEIKLIHRANQIVKNVPDNLPLDVLPVNSCGGIDRCSSBQ	859
Db	742	KLKAYTRYQLRGYTEDSQDLBEYLIRYNAKHETVNPQGT--GSLWPLSAPSPIGKCAHH-	798
Qy	860	YVDANLALENNGENGNMSSDSHAFSPHIDTGEIDLNENTGIWVYVFKIPTNGYATLGNLE	919
Db	799	-----SHHPSLDIDVCGTDLNEGLVWVYFKIQDGHARLGNLE	838
Qy	920	LVBEGPLSGETLEAQQOQOQWQDMARKRGASEKAYVAAQKAIIDLRFADYQDOKLNSGV	979
Db	839	FLEEKPLVGEALARKVRAEKWKDKREKLEWETINIVYKEAKESVDALFVNSQYDLQADT	898
Qy	980	EMSDMLAQNLVQSIPIVYNDALPEIDGMNYTSTELTNRLQQAOWNLYDLBNALPNGDPR	1039
Db	899	NIAMIHAADKRVSIREAYLPESVIPGVNAAIIFEELEGRIFTAFSLYDAENVTIKNGDPN	958
Qy	1040	NGLSDMNATSDVNY-QQLSDTSLVLIPIWNNSOVSOQFTVQPNRYVLRLVATARKGVGGDY	1098
Db	959	NGLSCMNVKGVHVBEGQNNHRSVLVPEWEAEVSEVQVRCBGYILRLVATYKGVGEGC	1018
Qy	1099	VIIRDGANQOTETLTFN-----ICDDDTGV-----	1122
Db	1019	VTIHEIENNTDELKFSNCVBEVYPNNVTVCNDYATATQEEYEGTYSRNRGVDCAYESNS	1078
Qy	1123	-LSADQTS-----YTKTVBEFTPTSTEQWIDMSE	1150
Db	1079	SVPADYASAEYKAYTDGRDRNPNCESNRGYDGYTFLPAGYVTKESLFFPETDKVWIEIGE	1138
Qy	1151	TEGVFNIESVELVLEE	1167
Db	1139	TEGTFIVDSVELLMKE	1155

RESULT 11
S02134
parasporal crystal protein cryIC1 - Bacillus thuringiensis (strain aizawai IC1)
N:Alternate names: delta-endotoxin IC1; entomocidal crystal protein
C:Species: Bacillus thuringiensis
A:Variety: strain aizawai IC1
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 01-Dec-2000
C:Accession: S02134; S04594
R:Haider, M.Z.; Ellar, D.J.
Nucleic Acids Res. 16, 10927, 1988
A:Title: Nucleotide sequence of a Bacillus thuringiensis aizawai IC1 entomocidal crystal
A:Reference number: S02134; MUID:89083518; PMID:3205732
A:Accession: S02134
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-1155 <HAI>
A:Cross-references: EMBL:X13233; NID:g40277; PIDN:CAA31620.1; PID:g40278
A:Experimental source: strain aizawai IC1
R:Haider, M.Z.; Ellar, D.J.

J. Mol. Biol. 208, 183-194, 1989

A:Title: Functional mapping of an entomocidal delta-endotoxin. Single amino acid change

A:Reference number: S04594; MUID:89382455; PMID:2789751

A:Accession: S04594

A:Molecule type: DNA

A:Residues: 429-449, 'A', 451-724 <HAM>

A:Cross-references: EMBL:X16315

A:Experimental source: strain aizawai ICI

C:Comment: This toxin is lethal to the larvae of lepidopterans and dipterans.

C:Superfamily: parasporal crystal protein

C:Keywords: delta-endotoxin

Query Match	34.08;	Score 2052.5;	DB 2;	Length 1155;
Best local Similarity	39.4%;	Pred. No. 6.3e-111;		
Matches	479;	Conservative 183;	Mismatches 388;	Indels 167; Gaps 29;
Qy	53	NPEL-----FCNPET-FISSSTWGTG---IGI---	VGQVLGALGVPPAGQIASVFSF	97
Db	4	NPINECIPYCNLSNPEVEVLGGRIHSTGTPDIDISLSLTQPLSEFPVPGVFLGLVDI	63	
Qy	98	IVGQLMPSSTVSVMIMIMKQVEDLIDOKITDSVRKKTALAGLQGLDGLDVYQKSLKNMLE	157	
Db	64	IWGIPIGPSQ---WDAPLVQIEQLINQRIEESPARNQAI SRL EGLSNLYQIVABSFREWEA	119	
Qy	158	NRMDTRARSVVVTVQYIALELDFVAKIPSPAI SGOEVP LLSVYQAQANLHLLLLDASIFG	217	
Db	120	DPTNPALREEMRIQFNDMNSALTAIPLFQAVQNYQVPLLSVYQAQANLHLSVLSDVSVFG	179	
Qy	218	AKWGFTPGELSTFVDROVTRTAQSDYCVKMYNTGLDKLAGTNAASVLYKHOFREMTLL	277	
Db	180	QRWGFDAAATINSRYNDLTRIGNYTDHAVRWYNTGLSERVWVGPSDROWIRTNQFRRELT	239	
Qy	278	VLDLVALFPNVDRTYPIETTAQLTRREYVYDPIVFNRETSGGFCRRWSLANSDISFSVEBS	337	
Db	240	VLDIVSLFPNVDRTYPIRTVSQLTRREIYTNVLENFD--GSP--RCSAQG-----IEG	289	
Qy	338	AVIRSPHLFDILSEIEFTYTRAGLPLNNTBYLYVWGHGSIK-----YENTNASSALERNY	392	
Db	290	S-IRSPHMLDILNSITITYT-----DAHRGEY--YWSGHQIMASVPVSGSPGFEFTPL---	Y 338	
Qy	393	GTITSNKIKYDILANKDIPQVRSIGADLANVYAAQVY--GVPIASFTLLDKNT-----GSGS	446	
Db	339	GTWGNAAPOQRIVAQLGGVYRTLSSTL---YRRFPNIGINNOQLSLVDGTETPAYGTSSN	395	
Qy	447	VGGFTYSKPHHTMQVCTQNTYNTIDEIPPENE--PLSRGYSGRHLSHITSY--SFSKSNASP	502	
Db	396	LPSAVYRKSGTV-----DSLDELPPQNNVPPRQGFPSHRLSHVSMTSGFSNVSVI	447	
Qy	503	ARYGNLPVPFAWTHRSADVTVVYSDKITQIPVKAHMTLVSGTTVIKPGPTGGMILKRTS	562	
Db	448	IR---PMPFSWIHRSAEPNNIIPSSQITQPLTKTSNLGSGTSSVWVGPGPTGGDILARTS	504	
Qy	563	SGPLAYTSVSKPSLQRYPARRYASTNNLRLFTVITSGTRIYSINVKNTWKNGDDLTPTN	622	
Db	505	PGQISTLRVNITAPLSQRYEVRIRYASTNNLQHLTSIDGRI INQENFSATMSSGNSGNLQSG	564	
Qy	623	TFDLATIGTAFTSNYSDSLTVGADSPASGEVYVDKFELIPVNATPEABEDLDVAKKAV	682	
Db	565	SPRIVGTPTTFPNSGSSVFTLSAHVPNSGNEVYIDRIEFPVPAEVTPEAEVDLEBAQKAV	624	
Qy	683	NGLPTSKKD--ALQTSVTDYQVQAANLVECLSDELYPNEKRMMLWDAVKEAKRLVOARNLL	741	
Db	625	NELFTSSNQIGLAKTDVTDYHIDQVSNLVECLSDPECLDEKKEKSEKVKHAKRLSDERNLL	684	
Qy	742	QDTGPNRING--ENGWTGSGTIEVAEGDVLFKDRLSLP LTSAREIDTETPTYLQQOIDES	799	
Db	685	QDNFPRGINQLDRGRGSGTDITIQGGDDVFKENYVTLGLTFD---ECYLTLYQKIDES	741	
Qy	800	LLKPYTRYKLKGFITSSQDLRLKILRRANQIVGNVPDNLPLDVLVPMVSCGIGDRCSEQ	859	
Db	742	KLKAYTRYLQURGYIBDSQDLRLIYLIRNACHETVNP-----GTGSLWLSAPS	790	
Qy	860	YVDANLALENNGENGNMSSDSHAFSPHIDTGEIDLNENTGCIWVVPKIPITTINGATLGNLE	919	

Db 791 PI-----GKAAHSHHSLDIDVCGTDLNEDLGWVIFKIKTQGHARLQWLE 838
Qy 920 LVEPGPLSGETLRAQOOQOQWQDMARKRGASEKAYVAAKQADRLPADYQDOKLNGV 979
Db 839 FLEEKPLVGEALARKVRAEKWKDKREKLEWETNIIVYKEAKESVDALPVNSOYDLQADT 898
Qy 980 EMSDLAAQNLVQSPYVYNDALPEIPGMNTYTFTELNRLOQAQWLYDLBNALPNDGPR 1039
Db 899 NIAMIHAADKRVHSIREAYLPESLIPGVNAAIFEELEGRIPTAPSLYDARNVINGDN 958
Qy 1040 NGLSDMNAATSDVNV-QQLSDTSVLVIPNNWNSOVQFTVQNYRVLVLTARKEGVGDY 1098
Db 959 NGLSCWNVKGVHDVEEQNHRSVLVPEWBAEVSQVPCRGYILRVATYKEGVGEC 1018
Qy 1099 VIIRGANQTTLTFFN-----ICDDTGV----- 1122
Db 1019 VTIIENNTDELKFSNCVEBEVYVNNVTVCNDYTATQEEYEGVTYSRNRGVYDGAYESNS 1078
Qy 1123 -LSADQTS-----YITKVEFTPESTEOVWIDMSE 1150
Db 1079 SVPADYASAYBEKAYTDCRRDNPBCSNRGYGDYPLPAGYVTKLEYFPETDKWVIBIG 1138
Qy 1151 TEGVFNIESVELLSE 1167
Db 1139 TEGTFIVDSVELLAMEE 1155

RESULT 12
JT0241
parasporal crystal protein - Bacillus thuringiensis (strain aizawai IPL7)
N:Alternate names: 135K insecticidal protein
C:Species: Bacillus thuringiensis
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 01-Dec-2000
C:Accession: JT0241
R:Shimizu, M.; Oshie, K.; Nakamura, K.; Takada, Y.; Oeda, K.; Ohkawa, H.
A:Title: Cloning and expression in Escherichia coli of the 135-kDa insecticidal protein
A:Reference number: JT0241
A:Accession: JT0241
A:Molecule type: DNA
A:Residues: 1-1176 <SHI>
A:Note: b. thuringiensis aizawai IPL7 produces similar 130K and 135K insecticidal proteins
C:Comment: The 135K protein has insecticidal activity against Plutella xylostella larvae
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.9%; Score 2048; DB 2; Length 1176;
Best Local Similarity 39.3%; Pred. No. 1.2e-110;
Matches 480; Conservative 174; Mismatches 414; Indels 152; Gaps 26;

Qy 53 NPEL-----FGMPET-FISSSTVQVG---IGI---VGQVLGALGVPPAGQIASFYSP 97
Db 4 NPINCEIPYNCNLSNPVEVLGGERIETGTPIDISLSTPQLLSEFVPGAGFVLGLVDI 63
Qy 98 IVGQLWPSTVSVMKIMQVEDLIDOKITDSVRTALAGLOGDGLDVYKSLKWL 157
Db 64 IWGIFGPSQ----WDAFLVQIEQLNQRIBEFARNOAISRLGLESLNLYOIAVESPREWEA 119
Qy 158 NRNDTRARSVVVTVIALELDPVAKIPGFAISQGVPLLSVYVAQAANLHLLLRDASIFG 217
Db 120 DPTNPALREEMRIQPNWNSALTTPLEFAVQNYQVPLLSVYVAQAANLHLSVLVDVSFVG 179
Qy 218 AEWGTPGEISTFYDQVTRTAQYSDYCVKMYNTGLDKLKGTAASWLKTYQPRREMTLL 277
Db 180 ORWGPDAAATINSRYNDLTELIGNYTDYAVRWYNTGLERVMGSDRDWRYNQFRELTLT 239
Qy 278 VLIDLALPNDYDTRTYETTAQLTRVYTDPIVENRETS-GGFCRRWSLNSDISFSVE 336
Db 240 VLIDVALPNSDSRYRTVTSQLRBYITNPVLENFDGSGFMGAQRIEQN----- 290
Qy 337 SAVIRSPHLFDLISIEFT-TRAGLPLNNTLEYWVGHSTK-----YQNTNASSALER 390

291 ---IRQPHLMDILNSITITVDHVG-----PNYWSGHQITASPVGSGPEFAPPLFG 339
Qy 391 NYGIIITSKIKYDYLANDKDIFQVRS-----IGADLANYYAQTGYVPYASFTLLDKRT 442
Db 340 MAGNAAPVVL---VSLTGLGIGIFRTLSPLRYRIILSGPNNQELFVLDTGTFEFSALTTWL 397
Qy 443 GSGSVGGFTYKSPHTTMQVCTQNTYNTIDBIPEPNE-PLSRGYSHRLSHITSYSPSKNAS 500
Db 398 PS-----TIYRQGTGTV-----DSLVDIIPQDNSVPPRAGFSHRLSHVTWLSQAAGAV 444
Qy 501 SPARYGNLPVPAWTHRSADVTNTVYSDKITQIPVVKAHTLVSGTTVIKPGPGTGGNILKR 560
Db 445 YTLR---APTFSWQHSRAEFNNIIPSSQITQPLTKSTNLGSGTSVWKGPGTGGDILRR 501
Qy 561 TSSGPLATYSVSKPSLQRYRARIYASTTNLRPLVPTISGTRIVSINNVKNTMNGDDLT 620
Db 502 TSPQGISLIRVNIITAPLSQRYRIRIYASTTNLRPLVPTISGTRIVSINNVKNTMNGDDLT 561
Qy 621 FNTPLDLAGTAPTSNYSDSLITVGCADSPASGGEVYVDKFLIPVNAIPEAEEDLDVAKK 680
Db 562 SGRFATVGTPTFPFNSGSSVFTLSAHVFNSEVYIDRIEVPVAVTFEASVDLERACK 621
Qy 681 AVNGLFTSKD-ALQTSVTDYQVQAANLVCLSDDELYPNEKRMVMDAVKRAKRLVQARN 739
Db 622 AVNELFTSSNOIGLKTQVTDYHIDQVSNLVECLSDDEPCLDEKQELSEKVKHAKRLSDERN 681
Qy 740 LLODTGPNRING--ENGWTSSTGIEVABGDVLFKDRSLRLTSAREIDTETPTLYVQOQID 797
Db 682 LLODPNFRGINRQLDRGWRGSTDITIQGGDDVFKENYVTLTGTFD---BCYPTLYLQKID 738
Qy 798 ESKLPYTRYKLGKPIGSSQDLLEIKLIRHRANQIVKQVDPN---LLPDLVPLVNSCGGID 853
Db 739 ESKLKAYTRYQLRGYIEDSQDLLEIYLRNAXHETVNVPGTCSLWPLSAQSPICKGEPN 798
Qy 854 RCBQOQYVDANLALENNGENG-NMSSDSHAFSHIDTGEIDLNTGNTGIWVVKIPPTNGY 912
Db 799 RCA--PHLEWNPDLDCSCRDGKCAHSHHFLSDIDVGCCTDLNEDLGWVIFKIKTQDGH 856
Qy 913 ATLGNLILVEGSPLSGETLRAQOOQOQWQDMARKRGASEKAYVAAKQADRLPADYQD 972
Db 857 ARLGNLFLERKPLVGEALARKVRAEKWKDKREKLEWETNIIVYKEAKESVDALPVNSQY 916
Qy 973 OKLNGSVEMSDMLAAQNLVQSPYVYNDALPEIPGMNTYTFTELNRLOQAQWLYDLRNL 1032
Db 917 DQLQADTIAMIHAADKRVHSIREAYLPESLIPGVNAAIFEELEGRIPTAPSLYDARNV 976
Qy 1033 IPNGDFRNLGSLDMNATSDVNV-QQLSDTSVLVIPNNWNSOVQFTVQNYRVLVLTAR 1091
Db 977 IKNGDFNNGLSQWNVKGVHDVEEQNHRSVLVPEWBAEVSQVPCRGYILRVATYK 1036
Qy 1092 EGVGDGVYIIRGANQTTLTFFN-----ICDD----- 1118
Db 1037 EGYGEGCVTHIENNTDELKFSNCVEBEVYVNNVTVCNDYTATQEEYEGVATSRNRGVN 1096
Qy 1119 DTGVLSDAQTS-----YITKVEFTPESTEOVWID 1147
Db 1097 EAPSPADYASVYBEKSYTDGRRNPENPCNRRGYDYTPLEPVGVYTKLEYFPETDKVWIE 1156
Qy 1148 MSETGVFNIESVELLSE 1167
Db 1157 IGTEGTFIVDSVELLAMEE 1176

RESULT 13
JC2219
parasporal crystal protein cryIIa - Bacillus thuringiensis
C:Species: Bacillus thuringiensis
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 01-Dec-2000
C:Accession: JC2219
R:Udayasuriyan, V.; Nakamura, A.; Mori, H.; Masaki, H.; Uozumi, T.
Biosci. Biotechnol. Biochem. 58, 830-835, 1994
A:Title: Cloning of a new cryII(a) gene from Bacillus thuringiensis strain FU-2-7 and a
A:Reference number: JC2219; MUID:94289859; PMID:7764972

Qy	913	ATLGNLELVEBGPGLSETLERAQQQOQWQDWARRRGSEKAYAAKQNIIDLRFADYQD	972
Db	857	ARLGNLEFLBEKPLVLGEALRAVGRABKQKREKLEWETNIYYKGAKEGSDALFVNSQY	916
Qy	973	OKLNSGVMSDMLAAQNLVQSIPVYNDAALPEIPGKNYTSFTLTLNRLOQAMNLYLRNA	1032
Db	917	DQLQADTNIAIHADKRVHSIREAVLPESLVIPGYNAAIFEELEGRISTAFSLYDARVV	976
Qy	1033	IPNGDFRNLGSDNNATSDNVNVOQLSDTSYIIPNNNSQVSQQFTYOPNRYVLRVTARK	1091
Db	977	IKNGDFPNNGLSCWNVKGVHDVBQNNQRSVLVVPVEAEVQSBRVRCPCRGYLLRTVAYK	1036
Qy	1092	BGVGDGVVIIRGANQTETLTFFN-----ICDD-----	1118
Db	1037	BGYGEGCVTIHEIENNTDLKFSNCVBEIYPNNVTVCNDYTVNQBEYGGAYTSNRGYN	1096
Qy	1119	DTGVLSADOTS-----YITTKVFETPSTEQWID	1147
Db	1097	EASVPADVASVVEEKSYYDGRRENCFENRGVRYDTPLPGVYVTKLEYFPETDKWIE	1156
Qy	1148	MSETGVFNIESVELVLEBE	1167
Db	1157	IGETGTFIVDSVELLMEBE	1176

RESULT 14

A41052

parasporal crystal protein cryAel - Bacillus thuringiensis (strain alesti)

C:Species: Bacillus thuringiensis

C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 01-Dec-2000

C:Accession: A41052

R:Lee, C.S.; Aronson, A.I.

J. Bacteriol. 173, 6635-6638, 1991

A:Title: Cloning and analysis of delta-endotoxin genes from Bacillus thuringiensis

A:Reference number: A41052; MUID:92011442; PMID:1655719

A:Accession: A41052

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1181 <LE>

A:Cross-references: GB:M65252; NID:G142874; PIDN:AAA22410.1; PID:G142875

C:Superfamily: parasporal crystal protein

C:Keywords: delta-endotoxin

Qy	53	NPEL-----FGNDET-FISSVTQTG---IGI---VQQLGALGVPPAGQIASFYSP	97
Db	4	NPKINECIPYCNLSNPEVEVLGGERIETGYTPIDISLSTQFLLSBFVPGAGFVLGLIDL	63
Qy	98	IVGQLPSSSTVSYEMIMKQVEDLIDQKITDSVRKLTALAGLOGLGDLGVYQKSLKNWLE	157
Db	64	IWGFVPGSQ---WDAPLVQIEQLISRIEFPARNQAIKRLGSLNLYQYIYALAFREWEA	119
Qy	158	NRMDTRARSVVVQYITALELDFVAKIPSPAISQGVPELLSVYQAAMHLHLLLRDASIFG	217
Db	120	DPTNPALREMRIQPDNDSALITAPLFTVQNYQVPLLSVYQAVNLHLSVLDRDVSFVG	179
Qy	218	AEWGFTPEGISFTFYDRQVTRTAQVSDYCVKMYNTGLDKLKGTTNAASWLKTHQFRREMTLL	277
Db	180	QRWGLDVATINSRYNDLTRLIGTYTDYAVRWYNTGLERVMGPDSDRWVRYNQFRRELTLT	239
Qy	278	VLDLVALFPNYDTPTPIETTAQLTREVTYTDPIFVNRETSGGCRRLSLNSDISFSVES	337
Db	240	VLDIVSLFPNYSRTKYPTIRTVSQUTREIYNVPLENFD--GSPRG-----SAQRIEQ	289
Qy	338	AVTRSPHLPDILGSIEPYTTRAGLPLNNTYLEYVUGHSIK-----YKNTNASSALERNY	392
Db	290	S-IRSPHLMILLINSITYTDAG-----GYTWSGHOIWSAPVSGSGPEFTPL---Y	338
Qy	393	GTITSNKIKYVDLANKDIFQVRSIGDALANYAQ--YGVGPVYASFLLDKNT---GSGS	446

Db 339 GTMGNAAPQQRIVAOAGQGVRTLS---STFRNPFIIGINNQRULSVLDGTEFAYGSSN 395
QY 447 VGGFTYSKPHTTMQVCTQNTYNTIDIPPE--NEPLSRGYSHRLSHITSY--SFSKNASSP 502
Db 396 LPSAVYRSGTV-----DSLDIIPQDNNVPRQGFSHLSHVSMFRSGFSNSVSI 447
QY 503 ARYGNLPVPAWTHRSADVNTVYSDKITQIPIVVKARTLVSGTIVTKGPGFTGNNILKRTS 562
Db 448 IR--APWFSMIHRSABFNIIIPSSQITQIPLTKSTNLGSGTSVVKPGFTGDDILRRTS 504
QY 563 SGPLATSVSVKSPLSQRYRARIYASTNLRLFTVTSIGTRYISVNVKNTKNGDGLTFN 622
Db 505 FQGISLTAVNTAPLSQRYRARIYASTNLRLFTVTSIGTRYISVNVKNTKNGDGLTFN 564
QY 623 TFDLATIGTAFPTFNSYSDSLTVGADSFASGEGVYVDFKPELIPVNAATFAEEDLDVAKKAV 682
Db 565 SFRVTGFTTTPNFNSGSSVFTLSAHVFNSEVYIDRIEFVPAEVTFAEYDLERAQAV 624
QY 683 NGLFTSKKD-ALQTSVTDYQVNAANLVECLSDLELYPNEKMLMDAVKAKRLVQARNLL 741
Db 625 NALFTSSNQIGLKTVDYHIDQVSNLVECLSDLELYPNEKMLMDAVKAKRLVQARNLL 684
QY 742 QDTGNRIN--GENGTSGTIEVAEGDVLKORSRLTSAREIDTETVPTLYQQIDBS 799
Db 685 QDPNFRGINRQDRGWRGSDTITIQGGDDVFKENYVTLPGTDP---ECYPTLYQKIDES 741
QY 800 LKPYTRYKLGKFGICSSODLEIKLRHRANQIVKQNP--DNLLPDLV--PVNSCGGIDRC 855
Db 742 KUKATRYQLRGYIIDSQDLEIYLRYNAKHETVAVPGTSLWPLSFSSICKCGEPNRC 801
QY 856 SBQYTSQDANLALENNGENG--NMSSDASHFSDHIDTGBIDLNTGIVVVKIPTTNGVAT 914
Db 802 A--PHLEWNPDLDCRDEKCAHSHHPSLDIDVGCIDLNBELGVMVIFKIKTQDGHAR 859
QY 915 LGNLELVREGPLSGETLRAQQOQOQWQKARKGASEKAYYAKQAIDRLFADYQDOK 974
Db 860 LGNLELVREGPLSGETLRAQQOQOQWQKARKGASEKAYYAKQAIDRLFADYQDOK 919
QY 975 LNSGVMSDLAAQMLVQSIPIVYNDALPEIICMNVYTSFELTNLQQAANLYDLRNAP 1034
Db 920 LQADTNIAHIHTADKRVHRIQAYLPESLVIQVNAIGIFEELEGRIFTAYSILYDARNV 979
QY 1035 NGDFRNLGSDNATSDVNV--QQLSDTSVLVI PNNSQVSVSQFTVQPNRYVLRTARKEG 1093
Db 980 NGDFRNLGSDNATSDVNV--QQLSDTSVLVI PNNSQVSVSQFTVQPNRYVLRTARKEG 1039
QY 1094 VGGGVYIIRDGANQITLTFTNIC-----DDD 1119
Db 1040 YGEGCVTTHIENNDDELKFSNCVEEVYPNNVTVCNEYTANQEBVGGAYTSCNRYDET 1099
QY 1120 TG---VLSADOTS-----YITKIVFTPTTEQVW 1145
Db 1100 YGSNYSVPADYASVVEEKAYTDGRENPCSNRGYGYDTPLPAGVYTKQLEYFPETDKVW 1159
QY 1146 IDMSFEGVNTIESVELVLEE 1167
Db 1160 IBIGETEGTFIVDSVELFLEE 1181

RESULT 15
A29125
parasporal crystal protein Bc2 - Bacillus thuringiensis subsp. kurstaki (strain HD-1)
C:Species: Bacillus thuringiensis subsp. kurstaki
C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 01-Dec-2000
C:Accession: A29125
R:Fischhoff, D.A.; Bowdish, K.S.; Perlak, F.J.; Marrone, P.G.; McCormick, S.M.; Niedermeyer, J.
A:Title: Insect tolerant transgenic tomato plants.
A:Reference number: A29125
A:Accession: A29125
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-1156 <FIS>

C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.8%; Score 2043; DB 2; Length 1156;
Best Local Similarity 39.1%; Pred. No. 2, 2e-110;
Matches 476; Conservative 185; Mismatches 389; Indels 168; Gaps 30;

QY 53 NPEL-----FGNPET-FISSSTVQTG---IGI-----VGQVLGALGVPPAQOIAFYSP 97
Db 4 NPINECIPYNCLSPREVEVLGGERIETGYTPIDISLSTQFLLSFVFGAGFVLGLVDI 63
QY 98 IVGQLWPSSTVSWEMIMKQVEDLDDKIDTSVRKTALAGLOGLDGLDVOYKSLKNWLE 157
Db 64 IWGIFGPGQ---WDAFLVQIEQLNQRIEERPAQNQALSRLEGLNLYQIYAESFWEA 119
QY 158 NRNDTRARSVVVTVYIALELDFVAKIPSPAISQGVPLLSVYQAANLHLLLRDASIFG 217
Db 120 DPTPALREEMRIQPNDMNSALTAIPLFAVQNYQVPLLSVYQAANLHLSVLRDVSFVG 179
QY 218 AEWGTPGEBISTFYDQRTVTRTAQYSDYCVKYNVTGLDKLKTGNTAAASLWKVHPRREMTLL 277
Db 180 QRMGFDATINSRYNDLTRELIGNYTHAVRWNTGLERVMGPDSDRDWIRYQFRRLTIT 239
QY 278 VLDLVALPNTDTRTPPIETTAQLTREVVTDPIVFNRETSGGFCRRWSLNSDISPSEVES 337
Db 240 VLDIVSLFPNDYSRTPIRTVSQLTREIYTNPLENFD--GSF--RGSAGQ-----IEG 289
QY 338 AVIRSPHPLFDILSELEFYTTTRAGLPLNNTLEYLVVGHSHK-----YKVNASSALLERNY 392
Db 290 S-IRSPHMLDILNSITIT---DAHRGEY--YWSGHQIMASPVGFSGPEFTFPL---Y 338
QY 393 GTITSNKIKYDVLANKDIPQVRSGLADLANYYAQVY--GVVYASFTLLDKNT-----GSGS 446
Db 339 GTMGNAAPQQRIVAOAGQGVRTLSSTL---YRPPFNIGINNQQSLVLDGTEFAYGSSN 395
QY 447 VGGFTYSKPHTTMQVCTQNTYNTIDIPPE--NEPLSRGYSHRLSHITSY--SFSKNASSP 502
Db 396 LPSAVYRSGTV-----DSLDIIPQDNNVPRQGFSHLSHVSMFRSGFSNSVSI 447
QY 503 ARYGNLPVPAWTHRSADVNTVYSDKITQIPIVVKARTLVSGTIVTKGPGFTGNNILKRTS 562
Db 448 IR--APWFSMIHRSABFNIIIPSSQITQIPLTKSTNLGSGTSVVKPGFTGDDILRRTS 504
QY 563 SGPLATSVSVKSPLSQRYRARIYASTNLRLFTVTSIGTRYISVNVKNTKNGDGLTFN 622
Db 505 FQGISLTAVNTAPLSQRYRARIYASTNLRLFTVTSIGTRYISVNVKNTKNGDGLTFN 564
QY 623 TFDLATIGTAFPTFNSYSDSLTVGADSFASGEGVYVDFKPELIPVNAATFAEEDLDVAKKAV 682
Db 565 SFRVTGFTTTPNFNSGSSVFTLSAHVFNSEVYIDRIEFVPAEVTFAEYDLERAQAV 624
QY 683 NGLFTSKKD-ALQTSVTDYQVNAANLVECLSDLELYPNEKMLMDAVKAKRLVQARNLL 741
Db 625 NALFTSSNQIGLKTVDYHIDQVSNLVECLSDLELYPNEKMLMDAVKAKRLVQARNLL 684
QY 742 QDTGNRIN--GENGTSGTIEVAEGDVLKORSRLTSAREIDTETVPTLYQQIDBS 799
Db 685 QDPNFRGINRQDRGWRGSDTITIQGGDDVFKENYVTLPGTDP---ECYPTLYQKIDES 741
QY 800 LKPYTRYKLGKFGICSSODLEIKLRHRANQIVKQNP--DNLLPDLV--PVNSCGGIDRC 855
Db 742 KUKATRYQLRGYIIDSQDLEIYLRYNAKHETVAVPGTSLWPLSFSSICKCGEPNRC 801
QY 856 SBQYTSQDANLALENNGENG--NMSSDASHFSDHIDTGBIDLNTGIVVVKIPTTNGVAT 914
Db 802 A--PHLEWNPDLDCRDEKCAHSHHPSLDIDVGCIDLNBELGVMVIFKIKTQDGHAR 859
QY 915 LGNLELVREGPLSGETLRAQQOQOQWQKARKGASEKAYYAKQAIDRLFADYQDOK 974
Db 860 LGNLELVREGPLSGETLRAQQOQOQWQKARKGASEKAYYAKQAIDRLFADYQDOK 919
QY 975 LNSGVMSDLAAQMLVQSIPIVYNDALPEIICMNVYTSFELTNLQQAANLYDLRNAP 1034
Db 920 LQADTNIAHIHTADKRVHRIQAYLPESLVIQVNAIGIFEELEGRIFTAYSILYDARNV 979
QY 1035 NGDFRNLGSDNATSDVNV--QQLSDTSVLVI PNNSQVSVSQFTVQPNRYVLRTARKEG 1093
Db 980 NGDFRNLGSDNATSDVNV--QQLSDTSVLVI PNNSQVSVSQFTVQPNRYVLRTARKEG 1039
QY 1094 VGGGVYIIRDGANQITLTFTNIC-----DDD 1119
Db 1040 YGEGCVTTHIENNDDELKFSNCVEEVYPNNVTVCNEYTANQEBVGGAYTSCNRYDET 1099
QY 1120 TG---VLSADOTS-----YITKIVFTPTTEQVW 1145
Db 1100 YGSNYSVPADYASVVEEKAYTDGRENPCSNRGYGYDTPLPAGVYTKQLEYFPETDKVW 1159
QY 1146 IDMSFEGVNTIESVELVLEE 1167
Db 1160 IBIGETEGTFIVDSVELFLEE 1181

Db 899 TNIAIHADKRVHSIREAYLPESLVIPGVNAALFEELEBGEIFTAPSLYDARNVIKNGDP 958
Qy 1039 RNLSDWNATSDNVV-QOLSDTSVLVTPNWNNSQVSQQFTVOPNTRYVLRVLTARKEGVGDG 1097
Db 959 NGLSCWNVKGHDVVEBQNNHRSVLAVPEWEAEVSEQVRVCPGRGYILRVTYAYKEGYGEG 1018
Qy 1098 YVIRDGANCOTETLTEN-----ICDDDTGV----- 1122
Db 1019 CVTIHEIENNTDELKFSNCVBEVYPNNVTTCNDYTATQEEYEGYTTSRNRYDGAYESN 1078
Qy 1123 --LSADQTS-----YITKTVBFTPSTEQVWIDMS 1149
Db 1079 SSVFADYASAEKAYTDCRDNPCESNRGYDYTLPAGVYTKLELYFPETDKVWIEIG 1138
Qy 1150 ETEGVFNIESVELVLEE 1167
Db 1139 ETEGTFIVDSVELLIMEE 1156

Search completed: June 21, 2004, 10:13:10
Job time : 37 secs

Query Match 66.3%; Score 4007; DB 1; Length 1157;
Best Local Similarity 56.1%; Pred. NO. 1.1e-220;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;
QY 1 MSPNNQRYEILDASSTSVSNSVRYPLANDOTTLONNYKYDLRMSEGENPELFGNP 60


```
Qy 1 MSPNNQNEYELDASSSTSVSDNSVRYPLANDQTTTLQNMNYKDYLRMSGENPELFGNP 60
Db 1 MSPNNQNEYELDASSSTSVSDNSIRYPLANDQTTTLQNMNYKDYLRMSGENPELFGNP 60
Qy 61 ETRFSS-STVGTGIGVGVGLGALGVPFAGQIASFYFVGVGLMSPSSVSVVEMIMKQVE 119
Db 61 GTFISAQDAVGTGIDIVSTIISGLGIPVLGVEFVLSLGLLWPNSENENWQIIFMRVE 120
Qy 120 DLIDQKITDSVRKKTALAGLQGLDLYQKSLKNWLENRNDTRARSVVVVTQYIALELDF 179
Db 121 ELIDQKILDSVRSRAIADLANSRIAVEYQNALEDKRNKPHSTRSAAALVKERFCGNAEAIL 180
Qy 180 VAKIPSPAIQSQVPLSVTAQAANLHLLLRDASIFGAEWGFTPGELISTFYDQVTRTA 239
Db 181 RTNMGSPSQNYETPLPPTAQAASLHLVNRDVQIYGKEWGPQNDIDLFYKQSVSYTA 240
Qy 240 QYSDYCVKYNVTGLDKLGTNAASWLKYHOPRREMTLLVLDLVALFPNYDRTYPIETTA 299
Db 241 RYSDHCVMYVGNAGLKLKGTGAKQWVDYRPRRENVMVLQVALFPNYDARIYPLETNA 300
Qy 300 QLTREVTDPVFNRETSGGFCRRWSLN-SDI-----SPSEVESAVIRSPHLDIL 349
Db 301 ELTREITDPV-----GSVTCOSSTLISWYDMPAALPFSFTLEN-LLRKEPDPFTLL 352
Qy 350 SEIEFTYTRAGLPLNNT-BYLEYVWVHSIKYKNTNASSALERNYGTITSNKIKYDILANK 408
Db 353 QEIRMYTS--FRQNGTIEYVWGGORLTLSYIYGSSP--NKYSGLVLAGAEDIIIPVQGN 407
Qy 409 DIFQVRSIGDLANYAQAQVGPVYASPTLLDKNTGSGVGFTYKSHPTTMQVCTQYNT 468
Db 408 DIYEV--VMTYIGRTYNSLGLVNPVTP-YFSNNYTK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPPENEPLSRGSHRLSHITSYFSKNAASSPARYGNLVPFAMTHRSADVTNTVY 525
Db 455 IDSGEELTYEN--YQSVSHRVSYITSPFIKSTGTV--LGVPIFGWTHSSASRNNFIY 509
Qy 526 SKKITQIPVKAHTLVSGTTVKGPG-FTGNILKRTSSGPLAY---TSVSVKSLPLSORY 581
Db 510 ATKISQIPINKASRTSGGAVNMFQGLYNGGVPVKLSGSGQVNLNLRVATDAKG-ASORY 568
Qy 582 RARIRYASTTNLRFVTVS-----GTRIYSINNVKTMNGDDLTFNTPLATIGTA-P 633
Db 569 RIRIRYASDRAGKP--TISSRSPENPATYSIAIYNTMTNASTLSTYSPAYASGPNIL 626
Qy 634 TFSNYSDSLTVGADSFASGGVYVDKFLIPVNATFEAEEDLDVAKKAVNGLFTSKDAL 693
Db 627 GISSRSTPDISITKEAGANLYIDRIEFPVNTLFEAEEDLDVAKKAVNGLFTNEKDAL 686
Qy 694 QTSVTDYQVNOAANLVECLSDLYPNEKRLMDAVKBAKRLVQARNLLQDTGFNRINGEN 753
Db 687 QTSVTDYQVNOAANLIECLSDLYPNEKRLMDAVKBAKRLVQARNLLQDTGFNRINGEN 746
Qy 754 GWTGSGTGLEAGDVLFKDRSLRLTSAREIDETPTVLYQOIDEILLKPYTRYKLGFI 813
Db 747 GWTGSGTGLEVEGDVLFKDRSLRLTSAREIDETPTVLYQOIDEILLKPYTRYKLGFI 806
Qy 814 GSSODLETKLRRHANOIVKXVPDNLDPVLVPMNSCGGIDRCSEQOQYVDANLLENNGEN 873
Db 807 GSSODLETKLRRHANOIVKXVPDNLDPVLVPMNSCGGIDRCSEQOQYVDANLLENNGEN 866
Qy 874 GNMSSDSHAFSHFDITGEBIDLNENTGIWVVPKIPTTNGYATIGMLELVEEGLSGETLER 933
Db 867 GNMSSDSHAFSHFDITGEBIDLNENTGIWVVPKIPTTNGYATIGMLELVEEGLSGETLEW 926
Qy 934 AQOEEQWQDKARKGRASEKAYYAAKQAIIDRLPADYQDOKLNGSVENSDMLAAQNLVQS 993
Db 927 AQOEEQWQDKARKGRASEKAYYAAKQAIIDRLPADYQDOKLNGSVENSDMLAAQNLVQS 986
Qy 994 IPYVYNDALEPEIPGNNYTSFTELTNRLQQAANLYDLRNAIPNGDPRNGLSDWNATSDNVV 1053
Db 987 IPYVYNDALEPEIPGNNYTSFTELTNRLQQAANLYDLRNAIPNGDPRNGLSDWNATSDNVV 1046
```

```
Qy 1054 QQLSDTSVLVIPNNWSQVSOFTQPNRYVRLVTRARKEGVGDGYVIIRDGANOTETLTF 1113
Db 1047 QQLSDTSVLVIPNNWSQVSOFTQPNRYVRLVTRARKEGVGDGYVIIRDGANOTETLTF 1106
Qy 1114 NICDDDTGVLSDAQTSYITKTVPSTPSTEQWMDMSETEGVFNIESVBLVLEE 1167
Db 1107 NICDDDTGVLSTQDTSYITKTVEFTPTSTEQWMDMSETEGVFNIESVBLVLEE 1160

RESULT 3
C8BA BACUK
ID C8BA BACUK STANDARD; PRT; 1169 AA.
AC Q45705;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cry8Ba (insecticidal delta-endotoxin
DE CrvIIIIB(a)) (Crystalline entomocidal protoxin) (134 kDa crystal
DE protein).
GN CRy8BA OR CRvIIIIB(A) OR 50C(B).
OS Bacillus thuringiensis (subsp. kumamotoensis). Bacillus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae;
OX NCBI_TaxID=132267;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-18746 / P850C;
RA Michaels T.E., Fonceirrada L., Narva X.B.;
RT "Process for controlling scarab pests with Bacillus thuringiensis
RT isolates.";
RL Patent number W09315206, 05-AUG-1993.
CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS. ACTIVE ON VARIOUS SCARABEID BEETLES.
CC -1- DEVELOPMENTAL STAGE: the crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -1- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-6ib.ch).
CC -----
CC EMBL; U04365; AAA21118.1; --
CC HSSP; P07130; 1DLIC
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin_C.
CC InterPro; IPR005639; endotoxin_N.
CC InterPro; IPR008979; Gal bind_like.
CC Pfam; PF00555; endotoxin; 1.
CC Pfam; PF03944; endotoxin_C; 1.
CC Pfam; PF03945; endotoxin_N; 1.
CC Toxin; Sporulation.
CC SEQUENCE 1169 AA; 133543 MW; 22EFPQF5BD699909 CRC64;

Query Match 52.2%; Score 3152; DB 1; Length 1169;
Best Local Similarity 54.0%; Pred. No. 6,4e-172;
Matches 653; Conservative 179; Mismatches 294; Indels 84; Gaps 25;

Qy 1 MSPNNQNEYELDASSSTSVSDNSVRYPLANDQTTTLQNMNYKDYLRMSGENPELFGNP 60
Db 1 MSPNNQNEYELDASSSTSVSDNSNRYPFANETNALQNDYKDYLRMSAGNVSEYPCSP 60
Qy 61 ETRFIS-STVGTGIGVGVGLGALGVPFAGQIASFYFVGVGLMSPSSVSVVEMIMKQVE 119
Db 61 EVFLSEDAVAAADIVGKLITGLGVFPVGPVSLYITQIDILMPSKQSKQSEIFPMEQVE 120
Qy 120 DLIDQKITDSVRKKTALAGLQGLDLYQKSLKNWLENRNDTRARSVVVVTQYIALELDF 179
Db 121 ELIDQKILDSVRSRAIADLANSRIAVEYQNALEDKRNKPHSTRSAAALVKERFCGNAEAIL 180
```



```
Db 121 ELVNOQITTEPARNQALRLQGLGDSFNVYORSQNLMLADNRDNRNLSVVRQAQFALDLD 180
Qy 180 VAKIPSPAIQSQVPLLSVYAQAANLHLLLRDASIFGAEWGFTPGBISTFYDRQVTRTA 239
Db 181 VNAIPLFAVNGQVPLLSVYAQAANLHLLLRDASIFGAEWGFTPGBISTFYDRQVTRTA 240
Qy 240 QYSDYCVKWNVTGLDKLKGNAASWLKYHQFRREMTLLVLDVALFPNYTRTYPIETTA 299
Db 241 KYNYCETWNTGLDRLGRTNTESWLKYHQFRREMTLLVLDVALFPNYTRTYPIETTSNP 300
Qy 300 QLTREYVTDPIVNRFTSGFCRWSLNSDISSEVESAVIRSPHLPDIILSEIEFYTTTA 359
Db 301 QLTREYVTDPIVNRFPANVGLCRWGTNPNTSELENAPIRPHLPDLRLNSLTISNR- 359
Qy 360 GLPLNTEYLEVYVGHISIKYKNTNASSALERNYGTITSNIKY---YDLANKDIFQVRSI 416
Db 360 -FPV-SSNFMDYVSGHTLRSYLNSAVQSDSGVLTITTRATNPVGDGTNR-----IEST 413
Qy 417 GADLANYYAQQVGPYASFTL--LDKNTGSGVGGFTYKPHHTTMQVCTQNYNTIDRPP 474
Db 414 AVDFRSALIGYVNRASFVPGGLFNGTTSFANGG-----CRDLYDTNDELPL 461
Qy 475 ENRPLRGYS-HRLSHITSYSPKX-ASSPARVGNLPVPAWTHRSADVTTVVYSDKLTQI 532
Db 462 DE---STGSSTHRLSHVTFPSQTNQAGSIANAGSVPTVYVTRDVLNNTITPNRITOL 518
Qy 533 PVVKAHTLVGGTTVIGKPGFTGNNILKRTSGGLAYTSVSKPLSQRVYARIRYASTTN 592
Db 519 PLVVASAPVSGTTLVKGPGFTGGILRTTNGFTGLRVTVNSPLTQOYLRLRVFASTGN 578
Qy 593 LRLPVTISGRTIYSINVKTMKRGDDLTFTNP---DLATIG---TAPTFSNYSDSLTVA 646
Db 579 PSIRVLGGVSGIDVRLGTMNRQELTYESFFTRPTTTPGPNPPTFTQAQELITVNA 638
Qy 647 DSFASGGEVVDKPELLPVNATFEAREEDLVAKAVNGLPTSKDALQTSVTDYVQNOAA 706
Db 639 EGVSTGGEYVIDRIEIVPNVPAAREEEDLEAAKAVASLPTRTDGLQVNVTDYQVDOAA 698
Qy 707 NLVECLSDRLYPNEKRLMDVAKKRLVOARNLLQDTGNRRING--ENGWTTGSGTGEVA 764
Db 699 NLVSLCLSDQEGHDKMLRAVRAKRLSRERNLLQDPDFTNTSTENGKASNGVTIS 758
Qy 765 EGDVLPDRSLRTSAREIDETTYPTLYQOIDBSLLKPYTRYKLKGFIGSSQDLKLI 824
Db 759 EGGPFPGKRALQLASAR---ENYPTVIYQKVDASVLKPYTRYRLDGFVKSSQDLKLI 814
Qy 825 RHRANQIVKXVPDNLDPVLPVNSCGGIDRCSEQOYVDAMLALNNGENGNWSSDSHAFS 884
Db 815 HHKRVHLVXKVPDNLVSDTYSKSCSGINRCDEQHVQMDLDAEHHPMDCCRAAQTHEFS 874
Qy 885 FHIDTGEIDLINENTGIWVFKIPTTNGVATLGNLELVEBGLSGETTLERAAQOQOQWQDK 944
Db 875 SYINTGDNASVDGCIWVLKVRTTDTGATLGNLELVEBGLSGESLEREQDRWAKWNAE 934
Qy 945 MARKRGASAKYAAKQAIIDRLPADIYQOQKLSNGVEMSDMLAAQNLVQSIPYVYNDALPE 1004
Db 935 LGRKRAEIDRVYLAQAQAINHLFVDYQDQQLNPBGLABEINAEASNLVESISGVYSDTLQ 994
Qy 1005 IPGANYTSPFELTNRLQOANLVDLRNAIPNGDPRNGLSDWNATSDVNVQQLSDTSVLVI 1064
Db 995 IPGINYEITYELSDRLQAQASYLTSRAVQNGDFNSGLSDWNMTMDASVQOQGNMHLVL 1054
Qy 1065 PNWNSQVSOQPTVQPNRYVLRVTRARKEGVDGVYIRDGANOTETTLFNICDDDTGVLIS 1124
Db 1055 SHWDAQVSOQLRVNPNCKYVLRVTRARKVGGDGVYTRDGAHGETLTFNACDYDVNGTY 1114
Qy 1125 ADQTSYITKTVETPSTEQWIDKSETGVFNIESVELVLEE 1167
Db 1115 VNDNSYITEEVVFPETKMAVEVSESGSFYIDSIEPIETQE 1157
```

RESULT 5

```
C9EA_BACTA STANDARD; PRT; 1150 AA.
ID C9EA_BACTA Q9ZNL9;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cry9Aa (Insecticidal delta-endotoxin
DB CryIXB(a) (crystalline entomocidal protoxin) (130 kDa crystal
GN CRY9EA OR CRYIXE(A).
OS Bacillus thuringiensis (subsp. aizawai).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1433;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSK-10;
RA Midoh N., Oyama K.;
RT "Bacillus thuringiensis cry gene for insecticidal crystal protein.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases
CC -/- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -/- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -/- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -/- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.eib-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB011496; BAA34308.1; -
DR HSSP; P07130; IDLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C_1.
DR Pfam; PF03945; endotoxin_N_1.
DR Toxin; Sporulation.
KW Toxin; Sporulation.
SQ SEQUENCE 1150 AA; 129895 MW; 7D6AB93D6EDC97EB CRC64;

Query Match 46.7%; Score 2825.5; DB 1; Length 1150;
Best local similarity 48.6%; Pred. No. 2.6e-153;
Matches 574; Conservative 197; Mismatches 364; Indels 47; Gaps 17;

Qy 1 MSPNNNEYELTDASSSTSVSDNSVRYPLANDQTTTLQNNMYKDYLRMSEGE-----NP 54
Db 1 MNRNPNNEYELIDAPYCGCPSDDVVRPLASDPNNAQFNQMYKEYLYQYDGYTGSGLNP 60
Qy 55 ELFGNPTFTSSSTVQTGIGVQVGLGALGVPPAGQIASFYSPVIGQWLPSTSVSWEMI 114
Db 61 NLSINPR-----DVLQTGINIVGRILFGLVPPAGQLVTFYTFLLNQLWPTNDNAVWEAF 115
Qy 115 MKQVEDLIDQKITDSVRKTAGLQGLDGLDVQKSLKNLENENDTRASRVVVTQYIA 174
Db 116 MAQIEELIDQKISAQVVRNALDDLTGLHDHYERYEYLAALBEWLERPNGARA-NLVTRFEN 174
Qy 175 LELOFVAKIPSPFALS-QGE---VPLLSVYAQAANLHLLLRDASIFGAEWGFTPGBISTF 230
Db 175 LATAFVTRMPSFGTGPCSQORDAVALLTVYAQAANLHLLLRDASIFGAEWGFTPGBISTF 234
Qy 231 YDRQVTRTAQYSDYCVKWNVTGLDKLKGNAASWLKYHQFRREMTLLVLDVALFPNYDT 290
Db 235 PNAQOQETRIYTHNCHVETYNRGLEDVGRGTNTESWLKYHQFRREMTLLVLDVALFPNYV 294
Qy 291 RTYPIETTAQLTRVYTDPIVNRFTSGFCRWSLNSDISSEVESAVIRSPHLPDIILS 350
```

Db	295	ROYPNGANPOLTREIYDPIYVNPANQGI CRRGNNPNYNTFSELENAPTRPPLHFERLN	354
Qy	351	EIBPYTTRAGLPLNNTYELVWGHSTKYKNTWASSALERNYGTITSNKIKYYDLANKOI	410
Db	355	RLTISRNRYPATPNS-FLDYWSGHTILOSQHANNPTTYETSYGOITSN-TRLFNTTN-GA	411
Qy	411	FOVRSGLADLANYAQYGVYPYASFLLDKNTSGSGVGGFTYKPHTTMVQCTQNYNTID	470
Db	412	RAIDSRARNFCNLYANLYGV--SSLNIFP--TG-----VMSBITNAAANTCRQDLTTTE	460
Qy	471	EIPENEPISRGYSHRLSHLTSYSPSKNASP-ARYGNLPVFAWTHRSADVNTYSDXI	529
Db	461	ELEPLENNP-----NLLSHTYFLRFNTTQGGPLATLGFVTVYVMTREDEVFTNTITADRI	515
Qy	530	TQJPPVKAHTLVSGTWTIKGPGFTGGNILKARTSSGLAYTSVS VKSPLSQRYRARIRYAS	589
Db	516	TQJPPWKASELGGTTVVKGPGFTGGDILRRITDGGAVGTRIANVWAPLTOQYRILRYAS	575
Qy	590	TTN--LRLPVTYISGTRIYS INVNKTMKGGDILTPTFDLATIGTAPTFPSNYSDSLTVGAD	647
Db	576	TTSPVNNLPVNNSSAA--GPTLPSTMAQNGSLTYESPFTLEVTHTIRFQSQDFTTLRLNIP	632
Qy	648	SPASGGRVYVYDKFELI PVNATFBAEEDLDVAKAVNGLPTSKDALQTSYTDYQVNOOAN	707
Db	633	PSISGQBYVYDKLEI VIPNPTREABEEDLEDAKAVASLPTRTDGLQVNTDYQVOOAN	692
Qy	708	LVECLSDELYPNEKRMILWDVAKRAKLVQARNLLQDTGFNRING--ENGWGTSGTIGIAVE	765
Db	693	LVSCLSDEQGHDKMKMLLEAVRAAKRLSRRENLLQDPFNEINSTENGKASNGVTISE	752
Qy	766	GDVLFKDRSLRLTSARBIDTETPTYLYQQIDESLLKPYTRYKLGFGIGSSQDLKILIR	825
Db	753	GGPFKGRALQASAR---ENYPTYIVQKYVDASTLKPTRYRYKLDGFPWSSQDLIELIH	808
Qy	826	HRANOIVKNYDPNLLPDVLPVNSCGGIDRCSEQYVDANLALNNGENGWNSDSHAFSP	885
Db	809	HHKYHLVKNYDPNLVSDTYSDGSCGINRCBEOHQVDVQDAEPHDKDCCERAAQTHBESS	868
Qy	886	HIDTGETDLNENTGIWVPKIPPTNGYATLGNLEBVEGPLSGSETLBRACQEQOQWQDM	945
Db	869	YIHTGDLNASVQGIWVVLQVTRTIDGATYATLGNLEVEGPLSGESLEBERQDNAKWNEEV	928
Qy	946	ARKRGASEKAYAAQAIIDLFDYQDQKLSGVMSMDLAAQMLVQSIPIVYNDALPEI	1005
Db	929	GRKRAETDRYQDAKQAINHLFVDYQDQOLSPEVGMADIIDAQNLIIASISDVYSDAVLQI	988
Qy	1006	PGANYTSFTLITRLELOANWLYDLRNALPBGDFPRLSDNNATSDVNNQOLSOTSVLVIP	1065
Db	989	PGINENYMTLSNRLQOASLYLTSRNVVQNGDFNSGLDSWNNAITDTTAVQDGNHMFVLVS	1048
Qy	1066	NWNSQVSQOFTVQPNRYVRLRVTKARKEGVGDGYVIIRDGANOTELTAFNICDDDTGVLSA	1125
Db	1049	HWDQAQVSQOFTVQPNCKYVLEVTAKKVGNGDGYVTIQDGAHRETLIFNACDYDVNGTHV	1108
Qy	1126	DQTSYITKTVBETPSTEQWIDMSETGEVFNIESVELVLEE	1167
Db	1109	NDNSYITKELAVFYKTHMHWVESETGTFTFYDSIBETIOB	1150

RESULT 6

Accession	Protein	Organism	Length (aa)	PI	pI	pI _{2D}	Ref
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	1
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	2
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	3
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	4
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	5
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	6
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	7
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	8
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	9
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	10
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	11
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	12
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	13
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	14
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	15
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	16
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	17
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	18
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	19
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	20
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	21
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	22
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	23
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	24
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	25
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	26
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	27
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	28
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	29
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	30
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	31
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	32
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	33
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	34
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	35
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	36
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	37
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	38
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	39
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	40
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	41
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	42
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	43
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	44
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	45
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	46
Q9DPA	BACTP	<i>Escherichia coli</i>	1169	4.8	4.8	4.8	47
Q9DPA	BACTP						

```
Db 464 RNSTDDELPSLNPEGD-IIRNYSHRLSHITQYRFOATGSGSPSTVSANLPTCVWTHRDVQL 522
Qy 521 TNYVSDKITQIPVWKAHTLVSGTIVKGPFGTGGNLTGKTSGLSSGPLAYTYSVSKSPLSQR 580
Db 523 DNTITANQITQLPLVRAVELSSGATVVGKPGFTGGDVIKRTNTGGFCAIRSVTGPULTQR 582
Qy 581 YRARIYASTTNLRLFTVITSGTRIYSINVNKTWKNKGGDLTFTNFDLATIGTATFESNYS 640
Db 583 YRIRFYASTIDFPFVTRGGTINFRFRTRMNRGQESYESTVETFTFPNFTQSOD 642
Qy 641 SLTVGADSPAGSGEVYVDKPELIPVNAITPEABEDLDVAKKAV-NGLFTSKKDALQTSVTD 699
Db 643 IIRTSIQSLSGENGKVDLRIEIPVNPAREABEDLEAKKAAQNLFRTRDGLQVNVTD 702
Qy 700 YQVQANLVCELSDELYPNEKMLMDAVKAEKLVQARNLLOQTGNRRING--ENGWTG 757
Db 703 YQVQANLVCELSDELYPNEKMLMDAVKAEKLVQARNLLOQTGNRRING--ENGWTG 762
Qy 758 STGIEVAGDVLKDRSLRLTSAREIDTETPTVLYQOIDEKLVKTRVYKLGFGIGSSQ 817
Db 763 SNGVTISEGGPFGRALQASAR----ENPTIYQKVDASVLKPTIRYRLOGGFVXSSQ 818
Qy 818 DLBIKIRIRANQIVKQVNPDLPLVFNASCGGIDRCSEQQYVDANLALENNIGE-NGNM 876
Db 819 DLBIKIRIRANQIVKQVNPDLPLVFNASCGGIDRCSEQQYVDANLALENNIGE-NGNM 878
Qy 877 SSSSHAPFHDGCEIDLNNWGVKIPPTNGYATLGNLHVEGSPGSGETLEBAQQ 936
Db 879 AAOCTHEPSSVINTGDLNASVDQGVVWLVKTRVTDGYATLGNLHVEGSPGSGETLEBAQQ 938
Qy 937 QEQWQDKMKRKGASEKAYYAAQKQIDRLPADYQDOKLNSGVEMKSLAAQNLVQSIPI 996
Db 939 DNAKWNAELGRKABIDRVYLAQAQINHLFVDYQDQQLNPEIGLASINEASNLVESISG 998
Qy 997 VYNDALPEIQQMYNTSFTLNRLOQANLYDLNPAIPNGDFRNLGLSDNATSDVNVQQL 1056
Db 999 VYSDTLQIICINYEIYELSDRLOQASLYTSERAVQNGDFNSGLSDNATSDVNVQQL 1058
Qy 1057 SDTSVLVPMNNSQVSOQFTVQPNRYRVLVARTARKEGVGDGYIIRDGANOTETLTNLC 1116
Db 1059 GNMHFLVLSHWDQVSOQLRVNPKCVLVRVARKVGGDGYVTRDGAHQHQLTTFNAC 1118
Qy 1117 DDDTGVLSADQTSVITKYTFPTSPTEQVMDMSSETGVNIESVELVLEE 1167
Db 1119 DYDVNGTVYNDNSYITBEVVFYPTGMMVEVSESGFYDISEFTETQ 1169
```

RESULT 7

```
CLKA_BACTM STANDARD; PRT; 1215 AA.
AC Q45715;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cryIka (Insecticidal delta-endotoxin
DE CryIka(a)) (crystalline entomocidal protoxin) (137 kDa crystal protein).
GN CRYIKA OR CRYIKA(a) OR CRYIka.
OS Bacillus thuringiensis (subsp. morrisoni).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1441;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=F190;
RX MEDLINE=96102856; PubMed=8586263;
RA Koo B.T., Park S.-H., Choi S.-K., Shin B.S., Kim J.I., Yu J.H.;
RT "Cloning of a novel crystal protein gene cryIka from Bacillus
RT thuringiensis subsp. morrisoni."
RL FEMS Microbiol. Lett. 134:159-164(1995).
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS. SELECTIVELY TOXIC TO ARTOGEDIA RAPAE
CC AND NOT ACTIVE ON PLUTELLA XYLOSTELLA.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
```

```
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC ENBL; U28801; AAB00376.1; -.
DR HSP; P02965; ICY.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
DR Toxin; Sporulation.
SQ SEQUENCE 1215 AA; 137378 MW; 7CAF7A3311893D9B CRC64;
```

Query Match 44.9%; Score 2714; DB 1; Length 1215;

Best Local Similarity 46.3%; Pred. No. 6.4e-147; Indels 130; Gaps 23;

Matches 581; Conservative 194; Mismatches 351;

```
Qy 1 MSPNQNEYELDASSSTVSNSVRYPLANDQTTTLQNMNYKDYLRMSEGENPELFGNP 60
Db 1 MNSRKNENEIINALSIPAVSNHSAQMDLSPD-----ARIEDSLCAEG-----NNI 47
Qy 61 ETFISSVTQGTIGVQVLTGALGVPFAGQIAGSYFIVGQLWPSSTVSVMEMKQVED 120
Db 48 DPFVASTVTGTISAGRIKLVGLVGPAGQIAGSYFIVGQLWPSSTVSVMEMKQVED 106
Qy 121 LI-DOKITDSVRKTAGLQGLDVLVYQKSLKNLNRNDTRARSVVVTVYIALELD 179
Db 107 IVROQITDSVRDPAIARLEGLGRYSYQALSTLDRNDARSIRIRYIALELD 166
Qy 180 VAKIPSPAIQGVPELVSVAAQANLHLLLRDASIFGAEWGFTFGEISTFYDQVTRTA 239
Db 167 TTAIPLESIRNEEVLVWVAQANLHLLLRDASIFGAEWGFTFGEISTFYDQVTRTA 226
Qy 240 QYSDYCVQWNTGDKLKGNAASWLKHQPRREMTLLVLDLVALFPNPDRTYPIETA 299
Db 227 EYSNHCQWNTGDKLKGNAASWLKHQPRREMTLLVLDLVALFPNPDRTYPIETA 286
Qy 300 QLTREVVYTD--IVPNRETSGGFCRRWSLNSDISFSVESAVIRSPHLPDILSEFTYT 357
Db 287 QLTREVVYTD--IVPNRETSGGFCRRWSLNSDISFSVESAVIRSPHLPDILSEFTYT 339
Qy 358 RAGLPLNNTLEYLVWVGHISIKYKNTNASSALERNYITSNKIYYDL--ANKDIFQVRS 415
Db 340 RSQV--GTTIMNLWAGHRITFNRIQGGSTSEMYGALT-NPVSVDIPFVNRDVYRTVS 395
Qy 416 LGADLANYYAQVGVYASFTLLDKNTGSGVGGFTSKPHTHQVCTQNTYDIEIPPE 475
Db 396 LAGGLGSLGIRYGLTRVDFMDIFRN-HPDITVTGLFHPGHA--GIATQVKDSDTELPE 452
Qy 476 --NEPLSRGYSHRLSHITSYSFSKNASSPARYGNLVPFWTHRSADVNTVYSDKITQIP 533
Db 453 TTEQPNYRAFSHLLSHI-----SMGPTQDVPVYVSWTHQSADRNTNTNSDRITQIP 504
Qy 534 VKAHTLVSGTIVKPGFTGGNLTGKTSGLSSGPLAYTYSVSKSPLSQRYARIRYSTNL 593
Db 505 LVKHTLVSGTIVKPGFTGGNLTGKTSGLSSGPLAYTYSVSKSPLSQRYARIRYSTNL 564
Qy 594 RLFVTISGTRIYSINVNKTWKNKGGDLTFTNFDLATIGTATFESNYS 653
Db 565 RIYTVAGERIPAGQPKTMDAGAPLTFQSPSYATINTAFTFPRSSSLTIGADTFSSGN 624
```

```
Qy 654 EYVYDFKELIPVATFEABEDLDVAKAVNGLFSTSKD-ALQTSVTVYQVNAANVECL 712
Db 625 EYVYDFKELIQVATFEABEDLDVAKAVNGLFSTSKD-ALQTSVTVYQVNAANVECL 684
Qy 713 SDLYPNKMLMDAVKAEKLVQARNLQDTGNRING--ENGWTSSTGLVAGGVLP 770
Db 685 SDFCLDKKELLBEVYKAKLSDEKLLQDPTTSISGQDRGWIGSTGISQGGDDIP 744
Qy 771 KDRSLRLTSAREIDTETPTLYQIDESLKPTRYKLGKFGISSDLEKLRHRANQ 830
Db 745 KENYVRLPGT--VD-ECYPTLYQIDESQLKSTRYQLRGVIEDSDLEKLRHRANQ 801
Qy 831 IVKRVNPLP-----DVLVNSCGIDRCSSQYVDANLALNNGENGM-SDSHAPSP 885
Db 802 ETLSPGTEPSSGSSGVYPSGRCGEPRNRCAPR--IENPDLDCSCRYGKCVHSHHPSL 859
Qy 886 HIDEGETDLNENTGIVWVEKIPPTNGVATLGNLVEGSLGETLERAQOQOQODKM 945
Db 860 DIDVGCETDLNEDLGVWIFKIQDGHAKLGNLEFIEBKPLGLKALGRVKAEEKWRDKY 919
Qy 946 ARKGAKEAYAAKQADRLFADYQDOKLNSGVEMSDMLAAQNLVOSIPVYVNDALPEI 1005
Db 920 EKQLEKRVYTEAKESVDALFVDSQYDKLOANTNIGIHCADKQVRIREPYLSELPVI 979
Qy 1006 PGMYTSTELNRLQOANWLYDLRNAIPNGDFRNLSDMNATSVNVQQLSDTSVLVLP 1065
Db 980 PSINAAIFEELEGGHIFRAYSIDYDARVINKGDFNGLSCVNVKGVDPVQVNEHRSVLVS 1039
Qy 1066 NNSQVSOQFTVQVNYRVLRVLTARKEGVGGVYIIRGQNTETLTFN----- 1114
Db 1040 EWEAEVSKVRVCPDRGYLLRVATYKGEYGGCVTHIEFEDNTDVLFRNVEBEVYVNN 1099
Qy 1115 --ICDDDTGVLSADOTS----- 1129
Db 1100 TVTCDNTTQASGSDTACNSYRGYEDGYENRYEPNPSAPVNYTPYEBGVTDTQGY 1159
Qy 1130 -----YIKTVETPSTPQWIDMSSETGVNIESVELVLEE 1167
Db 1160 NHCVSDRGYRNHTPLPAGYVTLLEYEYPTEPQWIEIGETGTFIVGSVELLMEE 1215

RESULT 8
C1BB BACTU STANDARD; PRT; 1229 AA.
AC Q45739;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Pesticidal crystal protein cryIbB (insecticidal delta-endotoxin
GN CRYIbB) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
OS Bacillus thuringiensis
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NREL B-21110 / BG5847;
RA Donovan W.P., Tan Y., Jany C.S., Gonzalez J.M. Jr.;
RT "Bacillus thuringiensis cryet4 and cryet5 toxin genes and proteins
RT toxic to lepidopteran insects.;"
RL Patent number US5322687, 21-JUN-1994.
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L32020; AAA2344.1; -
CC HSPF; P02965; ICYI.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin_C.
CC InterPro; IPR005639; endotoxin_N.
CC InterPro; IPR008979; Gal_bind_Like.
CC Pfam; PF00555; endotoxin; 1.
CC Pfam; PF03944; endotoxin_C; 1.
CC Pfam; PF03945; endotoxin_N; 1.
CC Toxin; Sporulation.
CC KW
CC SEQUENCE 1229 AA; 139769 MW; A4C949DB675C3269 CRC64;

Query Match 42.2%; Score 2552; DB 1; Length 1229;
Best Local Similarity 43.2%; Pred. No. 1.1e-137;
Matches 548; Conservative 205; Mismatches 375; Indels 140; Gaps 22;

Qy 1 MSPNOMEYEILDASSSTVSNSVRYPLANDQTTLQNMVNYKDYLRMSEGENPELFGNP 60
Db 1 MTSRKNELEINALSPTVSNPTQNLSPD-----ARIESLCVAB-----VNNI 47
Qy 61 EFTISSVTQGTGIGVQVLGALGVPFAGQIATSPYSFIVGQLWPSSTVSVMHMKVED 120
Db 48 DPVFASTVQGTGINAGRIILGVLGVFPAGQIATSPYSFIVGQLWPSGR-DPWEIFLHVEQ 106
Qy 121 LIDQKITDVRKMTALAGLOGLDGVYQKSLKMLENDRTRARSVVVTVYALAELEDPV 180
Db 107 LIROQVTEENTNTAARLEGLRGYSYQQALETWLDNRNDRARSIIILERYVALELDIT 166
Qy 181 AKIPFAISGOEVLVSYAQAAMHLHLLLDASIFGAENGFPTGEISTFYDRQVTRTAQ 240
Db 167 TAILPLFRIEVEVPLLVAYQAANLHLLLDASIFGSENGWASSDVNQYQEQIRYTEE 226
Qy 241 YSDYCVKYNNTGDKLKGTTNAAGLKYHQPREMTLLVLDLVALPNTDTRTYPIETTAQ 300
Db 227 YSNHCQWYNTGLNLRGTNAESWLYNQFRDLTLGLVLDLVALPSSYDTRTYPIETSAQ 286
Qy 301 LTREYVTDPIVPRNETSGGFCRMSLNSDISFSEVESAVIRSPHLFDILSEFYETTRAG 360
Db 287 LTREIYTDPIGRTNAPSGPASTNFWNNAPSFAIEBAIFRPHLLDPBQLTYIS--AS 344
Qy 361 LPLANTYELRYWGHGSIKIKYK----NTNASSALERNYGTITSNKIKYDLANKDIFQVRS 416
Db 345 SRWSSTQHMVYVGHRLNFRPIGTLNTSTOGLTNTSTINPVLQF---TSRDVYRTEBN 401
Qy 417 GADLANYAQVYGVVYASFTLLDKRNTSGSGVGGTYSKPHTTMQVCTQNTYIDIRPE- 475
Db 402 AGTNILETTPVNGVPMARFNPINPQ-NIYERGAITYSQYQGVGI--QLFDSLETLPPE 458
Qy 476 -NEPLSRGYSHRLSHITTSYSKSNASSPARVGN--LPVPAWTHRSADVNTVYSDKITQ 531
Db 459 TERPNYSYSHRLSHIGLI-----IGNTLRAPIVSWTHRSADRTNIGNRIYQ 507
Qy 532 IPVKAHTLVSGTTVIKPGPFTGGNLRKTSGLLAYTSVSKPSLSQRYARIRYASTT 591
Db 508 IPLVKALNLHSGVTVVGPGFTGGDILLRRTWTGTFGDIRLNLNINVELSQRYRIRYASTT 567
Qy 592 NLRFLVPTISGRYIYSINNVKTMNKGDDLTFTWTFDLATIGTATFTSNYSOSLTVGADSPAS 651
Db 568 DLQFTRINGTWNIGNFRTMNRGDNLEYSFRTAGFTFPFNFLNAQSTFTLGAQSP-S 626
Qy 652 GGEVYVDFKELIPVATFEABEDLDVAKAVNGLFSTSKD-ALQTSVTVYQVNAANVECL 710
Db 627 NORVYIDRVEPFAEVTFEABEDLDVAKAVNGLFSTSKD-ALQTSVTVYQVNAANVECL 686
Qy 711 CLSDLEYPNKMLMDAVKAEKLVQARNLQDTGNRING----- 751
Db 687 CLSDLEYPNKMLMDAVKAEKLVQARNLQDTGNRING----- 746
```



```
Db 621 GNGEYVVDKILIPVATFEAYDYLERAQEAVALFTNPNRLKTDVTDVHIDQVSNLV 680
Qy 710 ECLSDLEYENKRMKMDVAKEAKRLVQARNLIQDTGFNRIN----- 750
Db 681 ACLSDPECLDKKELKVKYAKRLSDERNLIQDPNFTSINKQDPPISTNEQSNFTSIHE 740
Qy 751 -GENGWTSGLIEVAGDVLFDKRLSLTSARBIOTETPTLYQOIDSLLKPTRYKL 809
Db 741 QSEHGWGSENITIQEGNDVFKNFTVTLPGT---FNECYPTLYQKIGESSELKAYTRYQL 797
Qy 810 KGFISGSDLEIKLRRHRAQIVKNVP--DNLLPDLV--PVNSCGSIDRCSQQVVDANL 865
Db 798 RYIIESDQLEIYLIRAKHETLDPVGTESLWPLSVESPIGRCEPNRCA--PHFBNP 855
Qy 866 ALENGENG-NMSSDSHAFSFHIDTGEIDNENTGIWVFKIPTTNGVATLGNLELVEEG 924
Db 856 DLDSCRDGKCAHSHHPSLDIDVGCTDLHENLGVWVFKIKTQEGHARLGNLSPIBEK 915
Qy 925 PLSGTTLERAQOQDQWQKARKGSEKAYAAKQALDRLPADYQOQKLSGVEMSDM 984
Db 916 PLLGALSRLVKRAKKRDKRKLQLETKRVVTEAKAVDALFVDSQYDLRQADTNIGM 975
Qy 985 LAAQNLVOSIPVYNDALPEIPGMNYSFTLTNRLQAAWNLVDLRNAIPNGDPRNGLSD 1044
Db 976 HAADKLVRIRAYLSLSELPVPGVNAIPEELEGGHIITALSIDARNVWVGDFNNGLTC 1035
Qy 1045 WNATSDVAVVQLSDTSVLVIPNNSQVQFTVQPNRYRLVLTATARKGVDGYYIIRDG 1104
Db 1036 WNVKGVHDVQSHHSDLVIPWMAEVSQAVRCPGCGVILLAVTAYKEGVBGCVTHI 1095
Qy 1105 ANQETTLTF----- 1113
Db 1096 ENNTDELKPKNREBEVPTGTCTNDYTAHQGTAGCADCACNSRNAGYEDAYEVDTTASV 1155
Qy 1114 -----NICDDTGVLSAD--QTSVITKVEFTSTROVIMDMSETGV 1154
Db 1156 NYKPYEETTYDVRDHNCSYDRGVNYPVPVAGVTKELSYFPETDTVMIEIGETBK 1215
Qy 1155 FNIESVELVLEBE 1167
Db 1216 FIVDSVELLMEE 1228
[1]
RESULT 10
C1BC BACTM STANDARD; PRT; 1233 AA.
AC Q45774.
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pesticidal crystal protein cryIbC (insecticidal delta-endotoxin
GN CRYIbC OR CRYIbC) OR CRYIbC.
OS Bacillus thuringiensis (subsp. morrisoni).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1441;
[1]
SEQUENCE FROM N.A.
RA Bishop A.H., Bone E.J., Ellar D.J.;
RT "Cloning of novel Bacillus thuringiensis delta-endotoxin.";
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46442; CAA86568.1; -
DR HSPF; P02965; ICYI.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
KW Toxin; Sporulation.
SQ SEQUENCE 1233 AA; 140451 MW; 7318382413529F21 CRC64;
Query Match 41.2%; Score 2491; DB 1; Length 1233;
Best Local Similarity 42.1%; Pred. No. 3.4e-134;
Matches 536; Conservative 208; Mismatches 384; Indels 144; Gaps 20;
Qy 1 MSPNQNEYEILDASSTSVSDNSVRYPLANDQTTTLQNMNRYKDYLRMSEGENPELFGNP 60
Db 1 MTSNRKNEIINALSIPVSNPSTQMLSPD-----ARIEDSLCVAB-----VNNI 47
Qy 61 ETPSSSTVQIGIGIVGVVAGALGVPPAGQIASFYSFIVGQLWPSSTVSVWEMIKQVED 120
Db 48 DPFVSASTVQTGINTAGRIILGVLPVFPAGQASLFSFLVGLWPSGR-DPEWIFLSEHVEQ 106
Qy 121 LIDQKITPSVRTALAGLQGLDGVYQKSLKWLNRNDTRARSVVVTVYIALELDV 180
Db 107 LIQQVTTTNTAARLEGLRGVRSYQQALETWLDNNDARSIIILERYVALELDIT 166
Qy 181 AKIPSAISGQEVPLSVYAQAANLHLLLRDASIFGAEWGFTPGCBISTFYDRQVTRTAQ 240
Db 167 TAILPLIRNBEVPLLMVYAQAANLHLLLRDASLFGSEWGMASDVQVYQBIRYTEE 226
Qy 241 YSDYCVKQWNTGLDKLGTNAASWLKYHQPREMILLVLDLVALFPNYTRTYPIETTAQ 300
Db 227 YSNHCQWYNTGLNLRGTNAESWLRYNQPRDLTLGLVLDLVALFPSTYTRTYPIETSAQ 286
Qy 301 LTRVYVTDPIVFNRETSGPCRRWSLNSDISFSEVESAVIRSPHLFDLSEIEFYTRAG 360
Db 287 LTRBIYTDPIGRNAPSGFASNTWNNNAPSALERAIRFRPHLLDPFEQLTIYS--AS 344
Qy 361 LPLANTEYLEWVGHISIKYK-----NTNASSALERNYGTITSNKIKYDILANKDIFQVRL 416
Db 345 SWSSTQHMVWVGHRLNFRPIGGTLNTSTQGLTNTSINPVTLQF---TSRDVYRTESN 401
Qy 417 GADLANYAQQVGVVYASFTLLDRKNTGSGSVGGFTYKPHHTNQVCTQNTYIDEIPPE- 475
Db 402 ACTNILFTTPVNGVPMWAFNFINPQ-NIYERGATTSQPYQGVGI--QLFDSFETELPET 458
Qy 476 -NEPLSRGYSRHLSHITSYSPSKNASSPARVGN--LPVFAWTHRSADVNTVYSDKITQ 531
Db 459 TERPNYESYRHLSHIGLI-----IGNTLRAPVYSWTHRSADRTNIGPNRIIQ 507
Qy 532 IPVVKAHTLVSGTTVIKGPFTGGNLIKRTSSGLAYTSVSKVSPISQRYRARIYASIT 591
Db 508 IPLVKALNLHSGVTVVGGPGFTGGDILRRNTNGTGDIRLNININPLSQRYRARIYASIT 567
Qy 592 NURLFVTTISGTRIYSINVKNTKNGKDDLTFNTFDLATIGTATFSPYSLSLTVGADSPAS 651
Db 568 DLQFFTRINGTTVNTGNFSRTNMRGNDLEYSFRTAGFTFPNPLAQSTFTLGAQSP-S 626
Qy 652 GGEVYVVDKLELIPVNAIPEABEDLVAKVAVNGLFTSKD-ALQTSVTDYQVNAANLVE 710
Db 627 NOEYIDRVFVPAEVTFAEYDLERAQEAVALFTNPNRLKTDVTDVHIDQVSNMVA 686
Qy 711 CLSDLEYENKRMKMDVAKEAKRLVQARNLIQDTGFNRIN----- 751
Db 687 CLSDDFCLDKKELKVKYAKRLSDERNLIQDPNFTFISQQLSPASIDQSNFNPINEL 746
```


Db 909 TKRNKLAQMTTETQAIYTRAKQALDNLFAAQSHLKIQVTFABIAARKIVOSIREVY 968
Qy 999 NDALPEIFGMYNTSFTELNRLOQANWLYDLRNAIPNGDFRGLSDMNATSDVNVQQLSD 1058
Db 969 MSMLSVVGVNHPITFELSGRVQRAFLQYDVRNVVRNGRFLNGLSDMIVTSDVNVQBEHG 1028
Qy 1059 TSVLYVFNNSQVSOQFTVQVQNYRYLVLTARKGVGQGVYIIRGANGTETLTFNCD 1118
Db 1029 NNVLVLNMDAQVLRNKLQYDQGRVLRVTVARKIGIGEGYITITDEBCHTDLQAFAC-E 1087
Qy 1119 DTGVLSDAQTSYITKTVEFTSTEQWIDMETEGVFNESVLEVLLEE 1167
Db 1088 EIDASNAFISGYITIKLELFPDTEKVKHIEIGETGIFLVESIELFLLEE 1136

RESULT 14
C1BD_BACTZ STANDARD; PRT; 1231 AA.
AC Q9ZA25;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cryIbD (Insecticidal delta-endotoxin
DE CryIb(d)) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
GN CryIbD OR CryIbD OR CryIaI OR CryIaI.
OS Bacillus thuringiensis (subsp. wuhanensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=52024;
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN=HD-525;
RX MEDLINE=20153386; PubMed=10688690;
RA Kuo W.-S., Lin J.-H., Tzeng C.-C., Kao S.-S., Chak K.-F.;
RT "Cloning of two new cry genes from Bacillus thuringiensis subsp.
RT wuhanensis strain".
RL Curr. Microbiol. 40:227-232 (2000).
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF LEPIDOPTERAN LARVAE. TOXIC TO PLUTELLA
CC XYLOSTELLA.
CC -!- DEVELOPMENTAL STAGES: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U07026; AD10292.1; -.
DR HSSP; P02965; IC1Y.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin C; 1.
DR Pfam; PF03945; endotoxin N; 1.
KW Toxin; Sporulation.
SQ SEQUENCE 1231 AA; 139654 MW; 129A0371CDDBE52 CRC64;

Query Match 40.5%; Score 2445; DB 1; Length 1231;
Best Local Similarity 42.0%; Pred. No. 1.4e-131;
Matches 537; Conservative 205; Mismatches 377; Indels 160; Gaps 26;
Qy 1 MSPNQNEYTLDASSSTVSVDNSVRYVPLANDQTTLQNMVYKDYLRMSGEPNLPFGP 60
Db 1 MTSRKNENEIINALSTIPAVSNHSAQMDSLD-----ARIEDSLCIAEGNNI-----NP 49

Qy 61 BTFISSSTVQGIQVGVLTGALGVPPAGQIASFVSPIVQGLWPSSTVSVMIMKQVED 120
Db 50 --LVSASTVQGTGINIAGRIILGVLPAGQIASFVSPIVQGLWPSGR-DPMEIFLEHVQ 106
Qy 121 LIDOKITDSVRKTAALAGIQLGDLGVYQKSLKNWLENRNDTRARSVVVTVQVIALELDFV 180
Db 107 LIRQOVTEHTANTALRLEGLGRVRSYQQALETWLDNRNDARSIIILERYVALELDIT 166
Qy 181 AKIPSAISQGVFLLSYAAQANLHLLLRDASIFGAEWGPTPGCIISTFYDQVTRTAQ 240
Db 167 TAIPLERIRNEEVPLLAVYAQAANLHLLLRDASIFGSEWGMASDQVNYQEIRYTEE 226
Qy 241 YSDYCVKMYNTGLDKLKTNAASHKTHQPREMTLLVLDLVALPNVOTRTPIETRAQ 300
Db 227 YSNHCYQMYNTGLNLRGTNAESWLRYNQFREDTLTGLVLDLVALPSPYDTRTPINTSAQ 286
Qy 301 LTRVETDPIVFNRETSQGFRRWSLNSDISFSEVBSAVIRSPHLFDLSEIEFTYTRAG 360
Db 287 LTRIIYTDPIORTNAPSGFASTNFWNNAPSAIEAAIFRPHLLDPBQLTIYS--AS 344
Qy 361 LPLNTEYLEYVWGHESIKYK----NTNASSALERNYGTITSNKIKYDLANKDIFQVRS 416
Db 345 SRMSSTQHMVWVGHRLNFRPIGGTLNTSTQCLTNNTSINPVLQF---TSRDVYRTESN 401
Qy 417 GADLANYYAQVGVYASFTLLDKNTGSGVGGFTYKPHHTTMQVCTQNTYNTIDEIPPE- 475
Db 402 AGTNILEFTTPVNGVPMARFNFIPQ-NIYERGAITYSQYQGVGI--QLFDSSETLPPET 458
Qy 476 -NEPLSRGYSRHLSHITSYSFSKNASSPARYN--LPVFAWTHRSADVNTVVSQDKITQ 531
Db 459 TERPNYSYSRHLSHIGLI-----IGNTLRAPVYSWTHRSABRTNTIGPNRIITQ 507
Qy 532 IPVVKATLTVSGTTVIKPGPTGGNILKRTSSG-----PLATYSVSVKSPLSOR 580
Db 508 IPAVKGRFLNG-SVISGPGTGGDVRLNRNNGNIQNRGVIETVPIQTSTST-----R 560
Qy 581 YRAIRVASTTNRLFTVITISGTRIYSINVKTMKGGDLTFTNTPLATIGTFTPSNYSD 640
Db 561 YRVRYVASVTSIELNVNLGNSSIFNTPLATAASLDNLQSGDPGVYELNNAFTSATGN- 619
Qy 641 SLTVGADSPASGGGVYVDKFLIPVNAFPAEEDLDVAKAVNGVLFSTKKD-ALQTSVTD 699
Db 620 --IVGARFNSANAEVIDRFEFIPVATFPAEYDLERAQKAVNALFTSTNPRRLATDVT 677
Qy 700 YQVNAQANLVECLSDLEYENKRLMDAVKAKRLVQARNLLQDTGFRING----- 751
Db 678 YHIDQVSNMVACLSDPECLDEKRELFEKVYAKRLSDERNLLQDPNFTFISQLSFASID 737
Qy 752 -----ENGWTSQIEVAGDVLFKDRSLRLTSAREIDTETVPTLYQOIGDES 799
Db 738 QGSNFTSINELSEHGWSGSENVITQEGNDVFKENVYVLPGT---FNECTPNLYQKIGES 794
Qy 800 LLKPYTRYKLGKFGSSQDLLEIKLIRHANOIVKNVP--DNLLPDLV--PVNSCGGIDRC 855
Db 795 ELKATRYQLAGYIEDSQDLIELIIRYNAKHETLDVPGTDSIMPLSVKSPICRGCEPNRC 854
Qy 856 SEQQYVDANLALENGNG-NNSSDSHAPSPHIDTGEIDLNTGNTGIWVVKIPPTNGTAT 914
Db 855 A--PHFEMNPDLDCSRDGERCAHSHHFTILDIDVGCTDLHENLGVVVPFKIKTQEGVAR 912
Qy 915 LGNLELVEBGPISGTELRQAQOQOQODKMKRKGSEKAYAAKQAIIDLPAVQDQK 974
Db 913 LGNLEFTEBKPLIGEARSVRAEKQWDRKKEKQLETKRVYTEAKETVDALFVDSHYNR 972
Qy 975 LNSGVMSDMLAAQNLVQSIPIVYNDALPEIPGMNYTSFTELNRLOQANWLYDLRNAIP 1034
Db 973 LQADTNIGMHAADRLVRIHEAYLPELPFPPIPGINAVIFEELENRISTAPSLYDARNV 1032
Qy 1035 NGDFRNLGSDMNATSDVNVQQLSDTSVLVIPNNSQVSOQFTVQVQNYRYLVLTARKGV 1094
Db 1033 NGDPNNGLSQWNVKGVHDVQSHRSRLVPEWEAEVSAQVAVRVCPRGVILRVATYKEGY 1092

Search completed: June 21, 2004, 10:10:58
Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:10:21 ; Search time 90 Seconds
(without alignments)
3663.699 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQBYEILDASSTVS.....MSETEGVNIESVELVLEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6044	100.0	1167	5 AAU80281	Aau80281 Bacillus
2	4008	66.3	1157	2 AAR28900	Aar28900 Toxin 50C
3	4007	66.3	1157	2 AAR25997	Aar25997 Delta-end
4	4007	66.3	1157	2 AAR33768	Aar33768 Bt isolat
5	4007	66.3	1157	2 AAU06418	Aaw06418 Antiscara
6	3995	66.1	1157	2 AAR27343	Aar27343 B.thuring
7	3994	66.1	1157	2 AAR44208	Aar44208 Bacillus
8	3662.5	60.6	1149	2 AAR32354	Aar32354 Coleopter
9	3662.5	60.6	1149	2 AAR51692	Aar51692 B.thuring
10	3263.5	54.0	1210	5 AAU99256	Aau99256 Bacillus
11	3246.5	53.7	1206	5 AAU99255	Aau99255 Bacillus
12	3152	52.2	1169	2 AAU06417	Aaw06417 Antiscara
13	3128	51.8	1157	2 AAU84581	Aaw84581 Amino aci
14	3128	51.8	1157	2 AAU84584	Aaw84584 Amino aci
15	3126	51.7	1157	2 AAU84571	Aaw84571 Amino aci
16	3125	51.7	1157	2 AAU84591	Aaw84591 Amino aci
17	3125	51.7	1157	2 AAU84575	Aaw84575 Amino aci
18	3124	51.7	1157	2 AAU84586	Aaw84586 Amino aci
19	3124	51.7	1157	2 AAU84582	Aaw84582 Amino aci
20	3124	51.7	1157	2 AAU84593	Aaw84593 Amino aci
21	3124	51.7	1157	2 AAU84588	Aaw84588 Amino aci
22	3124	51.7	1157	2 AAU84578	Aaw84578 Amino aci
23	3124	51.7	1157	2 AAU84576	Aaw84576 Amino aci
24	3123	51.7	1157	2 AAR48678	Aar48678 Insectici
25	3123	51.7	1157	2 AAU84570	Aau84570 Amino aci

26	3123	51.7	1157	2 AAU84566	Aaw84566 Amino aci
27	3123	51.7	1157	2 AAU84569	Aaw84569 Amino aci
28	3123	51.7	1157	2 AAU84568	Aaw84568 Amino aci
29	3123	51.7	1157	4 AAB82178	Aab82178 Cry9C #1.
30	3122	51.7	1157	2 AAU84583	Aaw84583 Amino aci
31	3122	51.7	1157	2 AAU84592	Aaw84592 Amino aci
32	3121	51.6	1157	2 AAU84572	Aaw84572 Amino aci
33	3121	51.6	1157	2 AAU84577	Aaw84577 Amino aci
34	3120	51.6	1157	2 AAU84590	Aaw84590 Amino aci
35	3120	51.6	1157	2 AAU84579	Aaw84579 Amino aci
36	3119	51.6	1157	2 AAU84589	Aaw84589 Amino aci
37	3118	51.6	1157	2 AAU84587	Aaw84587 Amino aci
38	3117	51.6	1157	2 AAU84573	Aaw84573 Amino aci
39	3115	51.5	1157	2 AAU84567	Aaw84567 Amino aci
40	3115	51.5	1157	2 AAU84585	Aaw84585 Amino aci
41	3114	51.5	1157	2 AAU84580	Aaw84580 Amino aci
42	3109	51.4	1157	2 AAU84574	Aaw84574 Amino aci
43	3100.5	51.3	1156	2 AAU46857	Aaw46857 Bacillus
44	3100.5	51.3	1156	2 AAU24960	Aay24960 Bacillus
45	3100.5	51.3	1156	4 AAU02034	Aau02034 B. thurin

ALIGNMENTS

RESULT 1
AAU80281
ID AAU80281 standard; protein; 1167 AA.
XX
AC AAU80281;
XX
XX
DT 30-JUL-2002 (first entry)
XX
DB Bacillus thuringiensis insecticidal protein.
XX
KW Insecticide; transgenic; Coleoptera larvae.
XX
OS Bacillus thuringiensis.
XX
PN JP2002045186-A.
XX
PD 12-FEB-2002.
XX
PF 03-AUG-2000; 2000JP-00236140.
XX
PR 03-AUG-2000; 2000JP-00236140.
XX
PA (SDSB-) SDS BIOTECH CORP.
XX
XX WPI; 2002-356468/39.
XX
XX A protein having insecticidal activity, a DNA encoding said protein, and an agent and a method for preventing harmful organisms.
XX
PS Claim 1; Page 9-12; 19pp; Japanese.
XX
XX This invention relates to a crystalline protein comprising a fully defined sequence and the nucleotide sequence encoding this protein. The protein of the invention is an agent for preventing harmful organisms comprising Bacillus thuringiensis serovar galleriae SDS502, its mutant or a microbe transformed by a DNA encoding the protein. This microbe can be used to produce a protein containing the protein, or containing a protein having insecticidal activity produced by the SDS502, its mutant or a transformed microbe, a microbe which is transformed by using the above DNA and produces the above protein having insecticidal activity, a plant or a seed transformed by using the above DNA, and Bacillus thuringiensis serovar galleriae SDS502 producing a protein comprising and producing a protein showing insecticidal activity. The protein of the invention may have insecticidal activity. The agent is used for preventing Coleoptera larvae. This sequence represents the Bacillus thuringiensis insecticide protein of the invention

SQ Sequence 1167 AA;

```
Query Match      100.0%; Score 6044; DB 5; Length 1167;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTLLQNNYKDYLRMSGENPELFCNP 60
DB 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTLLQNNYKDYLRMSGENPELFCNP 60

QY 61 ETPFSSSTVQTGIGIVGVLGALGVPAGQIASFYSFIVGQLWPSSTVSVMEMIMKQVED 120
DB 61 ETPFSSSTVQTGIGIVGVLGALGVPAGQIASFYSFIVGQLWPSSTVSVMEMIMKQVED 120

QY 121 LIDQKITDSVRKTAAGLQGLDGLDVTYQKSLKWLKLENRNDTRARSVVVTQVIALLELDPV 180
DB 121 LIDQKITDSVRKTAAGLQGLDGLDVTYQKSLKWLKLENRNDTRARSVVVTQVIALLELDPV 180

QY 181 AKIPSPALSGQVPLLSVYQAANLHLLLRDASIFGAEMGPTGCEISTFYDQVTRTAQ 240
DB 181 AKIPSPALSGQVPLLSVYQAANLHLLLRDASIFGAEMGPTGCEISTFYDQVTRTAQ 240

QY 241 YSDYCVKMYNTGLDKLKGNTAASMLKYHQFREMLLLVLDLVALFPNYDTRTPYIETTAQ 300
DB 241 YSDYCVKMYNTGLDKLKGNTAASMLKYHQFREMLLLVLDLVALFPNYDTRTPYIETTAQ 300

QY 301 LTREYVTDPIVFNRTSGGFCRRWSLNSDISPSEVESAVIRSPHLPDILSBIEFTYTRAG 360
DB 301 LTREYVTDPIVFNRTSGGFCRRWSLNSDISPSEVESAVIRSPHLPDILSBIEFTYTRAG 360

QY 361 LPLNNTVEYVWGHISIKYQNTNASSALERNYGTITSNKIKYDOLANKDIPFVRSIGADL 420
DB 361 LPLNNTVEYVWGHISIKYQNTNASSALERNYGTITSNKIKYDOLANKDIPFVRSIGADL 420

QY 421 ANYAQAIVGVPYASFTLLDKNTGSGVGGFTYSKPHITTMQVCTQNYNTIDIEIPPENEPLS 480
DB 421 ANYAQAIVGVPYASFTLLDKNTGSGVGGFTYSKPHITTMQVCTQNYNTIDIEIPPENEPLS 480

QY 481 RGYSHRLSHITSYSPSKNASSPARYGNLPVPAWTHRSADVNTVYSDKLTQIPVYKAHTL 540
DB 481 RGYSHRLSHITSYSPSKNASSPARYGNLPVPAWTHRSADVNTVYSDKLTQIPVYKAHTL 540

QY 541 VSGTIVIKGPGTGGNILKRTSGGLAYTSVSVKPLSORYRARIYASTTNLRLPVITS 600
DB 541 VSGTIVIKGPGTGGNILKRTSGGLAYTSVSVKPLSORYRARIYASTTNLRLPVITS 600

QY 601 GTRIYSINVNKTNKGDDLTFTNTFDLATIGTFTSNYSDSLTVGADSPASGGEVYVDKF 660
DB 601 GTRIYSINVNKTNKGDDLTFTNTFDLATIGTFTSNYSDSLTVGADSPASGGEVYVDKF 660

QY 661 ELIPVATPEABEDLDVAKAVNGLFTSKDALQTSVTDYQVQNAANLVECLSDLYPNE 720
DB 661 ELIPVATPEABEDLDVAKAVNGLFTSKDALQTSVTDYQVQNAANLVECLSDLYPNE 720

QY 721 KEMLDVAKEARLQARNLQDTGPNR INGENGTGSGIERSVAGDVLFKDRSLRLTSA 780
DB 721 KEMLDVAKEARLQARNLQDTGPNR INGENGTGSGIERSVAGDVLFKDRSLRLTSA 780

QY 781 REIDTETPTLYIQIDSLKPYTRYKLGFIGSSQDLLEIKLIRHRANQIVKNVPDNL 840
DB 781 REIDTETPTLYIQIDSLKPYTRYKLGFIGSSQDLLEIKLIRHRANQIVKNVPDNL 840

QY 841 PVLVPSNCGGIDRCSQYQVDAANLALNNGENGNMSDHSAPSFHIDTGEIDLNENTGI 900
DB 841 PVLVPSNCGGIDRCSQYQVDAANLALNNGENGNMSDHSAPSFHIDTGEIDLNENTGI 900

QY 901 WVVFKIPTTNGVATLGNLELVEGPLSGTETLRAQQEQOQWODKWKARKEGASEKAYYAAK 960
DB 901 WVVFKIPTTNGVATLGNLELVEGPLSGTETLRAQQEQOQWODKWKARKEGASEKAYYAAK 960

QY 961 QAIIDLRFADYQDQKLSNGVMSDMLAAQNLVQSIPIVYNDALPEIPGMNNTSFTLTNRL 1020
DB 961 QAIIDLRFADYQDQKLSNGVMSDMLAAQNLVQSIPIVYNDALPEIPGMNNTSFTLTNRL 1020

QY 1021 QOANWLYDLRNLAI PNNGDFRNLGLSDWNATSDVNVQQLSDTSVLVLPNANSQVSOQFTVQPN 1080
DB 1021 QOANWLYDLRNLAI PNNGDFRNLGLSDWNATSDVNVQQLSDTSVLVLPNANSQVSOQFTVQPN 1080

QY 1081 YRYVLVTVARKEGVGDGVYIIRDCANQOTETLTFFNICDDDTGVLSDADQTSYIITKTVEFTPS 1140
DB 1081 YRYVLVTVARKEGVGDGVYIIRDCANQOTETLTFFNICDDDTGVLSDADQTSYIITKTVEFTPS 1140

QY 1141 TEQWIDMSTEGVNIESVELVLREE 1167
DB 1141 TEQWIDMSTEGVNIESVELVLREE 1167

RESULT 2
AAR28900
ID AAR28900 standard; protein; 1157 AA.
XX
AC AAR28900;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-APR-1993 (first entry)
XX
DE Toxin 50C.
XX
KW Endotoxin; acarides; pest; Two Spotted Spider; mite; phytophagus.
XX
OS Bacillus thuringiensis; kumamotoensis PS50C.
XX
PN W09219106-A1.
XX
PD 12-NOV-1992.
XX
PF 30-APR-1992; 92WO-US003546.
XX
PR 30-APR-1991; 91US-00693210.
PR 13-SEP-1991; 91US-00759248.
PR 30-SEP-1991; 91US-00768141.
XX
PA (MYCO ) MYCOGEN CORP.
XX
PI Payne JW, Cannon RJC, Bagley AL;
XX
WPI: 1992-398411/48.
DR N-PSDB; AAQ30821.
XX
PT New Bacillus thuringiensis isolates and toxins - used for controlling
acacrid pests of livestock, fowl, stored prods. and plants.
XX
PS Claim 16; Page 53 + 49-51; 62pp; English.
XX
CC Gene sequences encoding a toxin which is active against acarides and is
obtainable from B. thuringiensis isolates PS17a, PS17b, 33f2, PS2A1,
PS69D1, PS86A1 and PS50C are given in AAQ30803-07 and AAQ30820-21
respectively. The toxin is a delta-endotoxin active against acarid pests,
including the Two Spotted Spider mite. The isolates can be used against
non-phytophagus mites such as acarid pests of livestock, fowl and stored
prods. The genes can be cloned and used to transform other hosts, which
can be used to control mites, or in the case of transgenic plants, be
resistant to mites. (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 1157 AA;

Query Match      66.3%; Score 4008; DB 2; Length 1157;
Best Local Similarity 66.1%; Pred. No. 1.7e-255;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTLLQNNYKDYLRMSGENPELFCNP 60
DB 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTLLQNNYKDYLRMSGENPELFCNP 60

QY 61 ETPFSSSTVQTGIGIVGVLGALGVPAGQIASFYSFIVGQLWPSSTVSVMEMIMKQVED 120
DB 61 ETPFSSSTVQTGIGIVGVLGALGVPAGQIASFYSFIVGQLWPSSTVSVMEMIMKQVED 120
```



```
QY 241 YSDYCVKYNATGDKLKGNTAASMLKYHQFRREMTLLVLDLVALFPNYDTIRTYPIETTAQ 300
DB 241 YSDYCVKYNATGDKLKGNTAASMLKYHQFRREMTLLVLDLVALFPNYDTIRTYPIETTAQ 300
QY 301 LTRVYVTDPIVFNRETSGPCRRWSLNSDI9SFSEVESAVIRSPHLPDILSEIEFYTTTAAQ 360
DB 301 LTRVYVTDPIVFNRETSGPCRRWSLNSDI9SFSEVESAVIRSPHLPDILSEIEFYTTTAAQ 360
QY 361 LPLNTEYLEYVWGHISIKYKNT-NASSALERNYGTITSNKIKYVDLANKDIFOVRESIGAD 419
DB 361 LPLNTEYLEYVWGHISIKYKNT-NASSALERNYGTITSNKIKYVDLANKDIFOVRESIGAD 419
QY 420 LANTYAQYGVPIYASFTLLDKNTGSGVGGFTYSKPHPTTMQVCTQNTYNTIDEIPENB-P 478
DB 420 LANTYAQYGVPIYASFTLLDKNTGSGVGGFTYSKPHPTTMQVCTQNTYNTIDEIPENB-P 478
QY 479 LSRGYSRHLSHITSVSFSKNASSPARYGNLPUPAFWTHRSADVTNTVYSKDKITQIPVKAH 538
DB 479 LSRGYSRHLSHITSVSFSKNASSPARYGNLPUPAFWTHRSADVTNTVYSKDKITQIPVKAH 538
QY 539 TLVSGTTVIKPGFTGGMILKARTSCPLAYTSVSKSPLSQRYRARIYASTNLRLLPVT 598
DB 539 TLVSGTTVIKPGFTGGMILKARTSCPLAYTSVSKSPLSQRYRARIYASTNLRLLPVT 598
QY 599 ISGTIRIYSINVNKTKGDDLTFTNFTLTIATGTAFTFNSYSDSLTVGADSPASGGEVYVD 658
DB 599 ISGTIRIYSINVNKTKGDDLTFTNFTLTIATGTAFTFNSYSDSLTVGADSPASGGEVYVD 658
QY 718 KFEIIPVNAATEAREDDVAKAVNGLETSSKDALQTSVTVYVNOAANVECLSDLEYP 718
DB 718 KFEIIPVNAATEAREDDVAKAVNGLETSSKDALQTSVTVYVNOAANVECLSDLEYP 718
QY 719 NEKRLMDAVKBAKLVQARMLQDQTFNFRINGENGWGTSGTIEVAREGDLVFKDRLSLT 778
DB 719 NEKRLMDAVKBAKLVQARMLQDQTFNFRINGENGWGTSGTIEVAREGDLVFKDRLSLT 778
QY 779 SAREIDTETPTLYLQOQIDESLLKPYTRYKLGFTGSSQDLEIKLIRHANOIVKQVDPN 838
DB 779 SAREIDTETPTLYLQOQIDESLLKPYTRYKLGFTGSSQDLEIKLIRHANOIVKQVDPN 838
QY 839 LLPDLVFNWCGGIDRCSEQQYVDNATLALNENGNGNMSDSSHAPSFIHDGTDLNENT 898
DB 839 LLPDLVFNWCGGIDRCSEQQYVDNATLALNENGNGNMSDSSHAPSFIHDGTDLNENT 898
QY 899 GIWVFKLPTTNGVATLGNLELVEBPLSGETLERAOQOQOQWODKARKGASKEYA 958
DB 899 GIWVFKLPTTNGVATLGNLELVEBPLSGETLERAOQOQOQWODKARKGASKEYA 958
QY 959 AKQADRLPADYQOQKLSGVEMSDMLAAQNLVQSIPIVYNDALPEIFGMNYTSTELTN 1018
DB 959 AKQADRLPADYQOQKLSGVEMSDMLAAQNLVQSIPIVYNDALPEIFGMNYTSTELTN 1018
QY 1019 RLQQAWEIYDLNATPNGDFRNLSDNATSDVNVQQLSDTSVLVIPNNWSQVSQOFTVQ 1078
DB 1019 RLQQAWEIYDLNATPNGDFRNLSDNATSDVNVQQLSDTSVLVIPNNWSQVSQOFTVQ 1078
QY 1079 PNRYVLVATARKGEGVGYIIRGANTQETLTENICDDDTGLVSLADQTSYITKVEFT 1138
DB 1079 PNRYVLVATARKGEGVGYIIRGANTQETLTENICDDDTGLVSLADQTSYITKVEFT 1138
QY 1139 PSTEQVIMDSGTGVPNIRSVELVEEB 1167
DB 1139 PSTEQVIMDSGTGVPNIRSVELVEEB 1167
QY 1129 PVTQMWIEISETGTFVIESVELIVDVB 1157
DB 1129 PVTQMWIEISETGTFVIESVELIVDVB 1157
```

RESULT 4

AAR33768

ID AAR33768 standard; protein; 1157 AA.

XX

AC AAR33768;

XX

DT 25-MAR-2003 (revised)

```
DT 15-JUL-1993 (first entry)
XX Bt isolate Ps50C.
XX Lepidopteran-active; toxin; Ti; Ri; plant; cell; Bt.
XX Bacillus thuringiensis.
XX WO9304587-A1.
XX 18-MAR-1993.
XX 11-SEP-1992; 92MO-US007697.
XX 12-SEP-1991; 91US-00758020.
XX 08-SEP-1992; 92US-00941650.
XX (MYCO ) MYCOGEN CORP.
XX Uyeda KA, Bradfish GA;
XX WPI; 1993-100566/12.
XX N-PSDB; AAQ38653.
XX Controlling lepidopteran pests - using compen. of Bacillus thuringiensis
XX strains or plants or microorganisms transformed with their toxin genes.
XX Claim 12; Page 22-25; 38pp; English.
XX The sequences given in AAR33768-70 represent lepidopteran-active toxins.
XX The DNA encoding these sequences were used within a Ti or Ri plasmid, to
XX transform plant cells. Whole plants can then be regenerated from the
XX transformed cells. The toxin may also be produced by cloning Bacillus
XX thuringiensis (Bt). It may then be applied directly to the plant locus.
XX (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to
XX correct PR field.)
XX SQ Sequence 1157 AA;
Query Match 66.3%; Score 4007; DB 2; Length 1157;
Best Local Similarity 66.1%; Pred. No. 1.9e-255;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;
QY 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTLLQNNYKDYLRMSGGENPELFGNP 60
DB 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTLLQNNYKDYLRMSGGENPELFGNP 60
QY 61 ETEFSSSTVQGTGIGVGOVLGALGVPPRAGQIASFYSPVGOVLPSPSTVSVWEMIKQVED 120
DB 61 ETEFSSSTVQGTGIGVGOVLGALGVPPRAGQIASFYSPVGOVLPSPSTVSVWEMIKQVED 120
QY 61 ETEFSSSTVQGTGIGVGOVLGALGVPPRAGQIASFYSPVGOVLPSPSTVSVWEMIKQVED 120
DB 61 ETEFSSSTVQGTGIGVGOVLGALGVPPRAGQIASFYSPVGOVLPSPSTVSVWEMIKQVED 120
QY 121 LIDQKIDTSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVQVIALELDPV 180
DB 121 LIDQKIDTSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVQVIALELDPV 180
QY 181 AKIISPALSGQEVPLLSVYAAQANLHLILRLDASIFGAEWGTGCEISIFDYDQVTRTAQ 240
DB 181 AKIISPALSGQEVPLLSVYAAQANLHLILRLDASIFGAEWGTGCEISIFDYDQVTRTAQ 240
QY 241 YSDYCVKYNATGDKLKGNTAASMLKYHQFRREMTLLVLDLVALFPNYDTIRTYPIETTAQ 300
DB 241 YSDYCVKYNATGDKLKGNTAASMLKYHQFRREMTLLVLDLVALFPNYDTIRTYPIETTAQ 300
QY 301 LTRVYVTDPIVFNRETSGPCRRWSLNSDI9SFSEVESAVIRSPHLPDILSEIEFYTTTAAQ 360
DB 301 LTRVYVTDPIVFNRETSGPCRRWSLNSDI9SFSEVESAVIRSPHLPDILSEIEFYTTTAAQ 360
QY 361 LPLNTEYLEYVWGHISIKYKNT-NASSALERNYGTITSNKIKYVDLANKDIFOVRESIGAD 419
DB 361 LPLNTEYLEYVWGHISIKYKNT-NASSALERNYGTITSNKIKYVDLANKDIFOVRESIGAD 419
QY 420 LANTYAQYGVPIYASFTLLDKNTGSGVGGFTYSKPHPTTMQVCTQNTYNTIDEIPENB-P 478
DB 420 LANTYAQYGVPIYASFTLLDKNTGSGVGGFTYSKPHPTTMQVCTQNTYNTIDEIPENB-P 478
```


QY 779 SAREIDTETPTLYXQIDBSLLKPYTRYKLGKFGIGSSODLEIKLIHRANOIVKVPDN 838
Db 775 GAREIDTETPTLYXQIDBSLLKPYTRYKLGKFGIGSSODLEIKLIHRANOIVKVPDN 834
QY 839 LLPDLVLPVNSCGGIDRCSEQYVDANALENNGNGMSSDSHAFSPHIDTGRIDLNENT 898
Db 835 LLPDVSPVNSDGSINRCSEQYVNSRL-----EGENRSGDAHEFSLPIDIGELDYNEA 888
QY 899 GIWVFKIPTNGYATIGNLELVEEGPLSGETTERAQOQOQODKMARKRGASEKAYYA 958
Db 889 GIWVGKIDTPEGYATIGNLELVEEGPLSGDALERLQREBQOQKIOWTRREEDTDRYMT 948
QY 959 AKQAIIDLFPADYQODKLNKNGVMSDMLAAQNLVQSIDPYVNDALPEIPGANYTSFTBLTN 1018
Db 949 SKQAVRLYSYQDQQLNPVEITDLTAAQDLIQSIDPYVNEFPPEIPGANYTKFTBLTD 1008
QY 1019 RLQQAANLYDLRNLNPGDFRNLGSDWNATSDVNVQOLSDTSLVLPNNMSOVSOQFTVQ 1078
Db 1009 RLQQAANLYDORNLNPGDFRNLGSDWNATPGVEVQOQINETSVLVLPNNMSOVSOQFTVQ 1068
QY 1079 PNTRYVLRVTKARKEGVDGYVIRIDGANOTETLTFTNICDDDTGVLSDQTSYITKTVEFT 1138
Db 1069 PNQRYVLRVTKARKEGVDGYVIRIDGANOTETLTFTNICDDDTGVLSDQTSYITKTVEFT 1128
QY 1139 PSTEQWIDMSETGVPNIESVELVLEER 1167
Db 1129 PYTDQWIBISETEGTPYIESVELIVDVE 1157

RESULT 7
AA44208
ID AA44208 standard; protein; 1157 AA.
AC AA44208;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 09-DEC-1995 (first entry)
XX
DE Bacillus thuringiensis crystal protein PS50C.
KW Crystal protein; delta endotoxin; acaricide; pesticide;
KW biological control agent; transgenic plant; crop improvement.
XX
OS Bacillus thuringiensis serovar kumamotoensis; (isolate PS50C).
XX
XX US5262158-A.
XX
PD 16-NOV-1993.
XX
PP 30-APR-1992; 92US-00876280.
XX
PR 20-APR-1991; 91US-00693210.
PR 13-SEP-1991; 91US-00759248.
PR 30-SEP-1991; 91US-00768141.
XX
XX (MYCO) MYCOGEN CORP.
XX
XX Payne JM, Cannon RJC, Bagley AL;
XX
DR WPI; 1993-377387/47.
DR N-PSDB; AAQ51704.
XX
XX Controlling acaride pests e.g. two spotted spider mite - utilising
PT Bacillus thuringiensis delta-toxins cloned into microbe hosts.
XX
PS Disclosure; Col 71-76; 42pp; English.
XX
CC DNA encoding the insecticidal toxin can be cloned into baculo viruses and
CC transferred to other host microbes, preferably E. coli NM522(pMYC 2320)
CC NRRL B-18769, to control acaride pests, or to plants which become
CC resistant to the acaricide pests. Specifically, the two-spotted spider
CC mite (Tetranychus urticae) is controlled. (Updated on 25-MAR-2003 to

CC correct PF field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 1157 AA;
Query Match 66.18; Score 3994; DB 2; Length 1157;
Best Local Similarity 66.08; Pred. No. 1.4e-254; Mismatches 240; Indels 14; Gaps 7;
Matches 771; Conservative 144;
QY 1 MSPNNQNEYIELDASSSTVSVDNSRYRPLANDOTTTLQNNMYKYDYLKRMSEGENPELFGNP 60
Db 1 MSPNNQNEYIELDASSSTVSVDNSRYRPLANDOTTTLQNNMYKYDYLKRMSEGENPELFGNP 60
QY 61 ETPISSSTVGTGIVQVIGALGVPFAGQIASFYSPFIVGQWPSSTVSVMEMIMQVED 120
Db 61 ETPISSSTVGTGIVQVIGALGVPFAGQIASFYSPFIVGQWPSSTVSVMEMIMQVED 120
QY 121 LIDOKITDSVRKATAGLQGLGVYOKSLKNLNRNDTRARSVVVTVYALELDPV 180
Db 121 LIDOKITDSVRKATAGLQGLGVYOKSLKNLNRNDTRARSVVVTVYALELDPV 180
QY 181 AKIPSPFAISQOEVPFLSVYQAQAANLHLLLRDASIFGAEMGFTFGEISTFYDRQVTRTAQ 240
Db 181 SSIPSPFAVSGHEVLLAVYQAQAANLHLLLRDASIFGAEMGFTFGEISTFYDRQVTRTAQ 240
QY 241 YSDYCVQWYNTGLDKLGTNAASWLKTHQPRREMTLLVLDLVALFPNYDTRTYPIETTAQ 300
Db 241 YSDYCVQWYNTGLDKLGTNAASWLKTHQPRREMTLLVLDLVALFPNYDTRTYPIETTAQ 300
QY 301 LTRDVYDPIVFNRETSGGFCRRWSLNSDLSFVSBSAVIRSPHLPDLSIEFYVTRAG 360
Db 301 LTRDVYDPIVFNRETSGGFCRRWSLNSDLSFVSBSAVIRSPHLPDLSIEFYVTRAG 360
QY 361 LPLNTEYLEYVWGHISIKYKNT-NASSALERNYGTITSNKIKYDLANKOIFQVRSAGD 419
Db 361 ITLNDAIYINWGHITLYKERTADSTVYTYANYGRITSEK-NSPALEDRODIFEINSTVAN 419
QY 420 LANYAQVYQVYASFTLLDKNTGSGVGGFTYSKPHITMQVCTQNTNTIDEIIPENB-P 478
Db 420 LANYAQVYQVYASFTLLDKNTGSGVGGFTYSKPHITMQVCTQNTNTIDEIIPENB-P 478
QY 479 LSRGYSRHSHTSYSPSKNASSPARYGNLPVPAWTHRSADVTNTVYSDKITQIPVVKAH 538
Db 478 VASYSRHSHTSYSPSKNASSPARYGNLPVPAWTHRSADVTNTVYSDKITQIPVVKAH 538
QY 539 TLVSGTIVIKGPGFTGNIILKRTSSGPLAYTSVKSPLSQRYRARIYASTTMLRLEVT 598
Db 539 TLVSGTIVIKGPGFTGNIILKRTSSGPLAYTSVKSPLSQRYRARIYASTTMLRLEVT 598
QY 599 ISGTIYISINVNTKMKGGDLTFNTPLATIGTAFPTPSYSDSLTVGADSPASGEVYVD 658
Db 599 ISGTIYISINVNTKMKGGDLTFNTPLATIGTAFPTPSYSDSLTVGADSPASGEVYVD 658
QY 659 KPELI PUNATFEAREDDLVAKKAVNGLFTSKDALQTSVTDYQVNOANLVECLSDLYP 718
Db 659 KPELI PUNATFEAREDDLVAKKAVNGLFTSKDALQTSVTDYQVNOANLVECLSDLYP 718
QY 719 NEKRLMDAVKEARLVQARNLLQDTGFNRINENGWGTSGTIGIEVABGDVLPFORSRLT 778
Db 719 NEKRLMDAVKEARLVQARNLLQDTGFNRINENGWGTSGTIGIEVABGDVLPFORSRLT 778
QY 779 SAREIDTETPTLYXQIDBSLLKPYTRYKLGKFGIGSSODLEIKLIHRANOIVKVPDN 838
Db 775 GAREIDTETPTLYXQIDBSLLKPYTRYKLGKFGIGSSODLEIKLIHRANOIVKVPDN 834
QY 839 LLPDLVLPVNSCGGIDRCSEQYVDANALENNGNGMSSDSHAFSPHIDTGRIDLNENT 898
Db 835 LLPDVSPVNSDGSINRCSEQYVNSRL-----EGENRSGDAHEFSLPIDIGELDYNEA 888
QY 899 GIWVFKIPTNGYATIGNLELVEEGPLSGETTERAQOQOQODKMARKRGASEKAYYA 958
Db 889 GIWVGKIDTPEGYATIGNLELVEEGPLSGDALERLQREBQOQKIOWTRREEDTDRYMT 948
QY 959 AKQAIIDLFPADYQODKLNKNGVMSDMLAAQNLVQSIDPYVNDALPEIPGANYTSFTBLTN 1018
Db 949 SKQAVRLYSYQDQQLNPVEITDLTAAQDLIQSIDPYVNEFPPEIPGANYTKFTBLTD 1008

Db 949 SKQAVRLYADYQDLQLPDPVEITDLTAQDLIOQIPYVYNEMFPEIPGMNFKTELTD 1008
 Qy 1019 RLQQAOWNLYDLRNLTPNGDFRGLSDWNATSDVNVQQLSDTSVLVIPPNNNSQVSQFTVQ 1078
 Db 1009 RLQQAOWNLYDQNALPNGDFRGLSNWATPGVEVQQLNHHSVLVIPPNDQVSQFTVQ 1068
 Qy 1079 PNRYVLVLTARKEGCGVGYIIRDGANQTETLTNFCDDDTGVLASADQTSYITKTVFT 1138
 Db 1069 PNQVRLVLTARKEGCGVGYIIRDGNGQSETLTFTSASDYDTNGVYNDQGTGVIITKTVFT 1128
 Qy 1139 PSTQVWIDMSETEGVFNIESVELVLEE 1167
 Db 1129 PYTQMWIEISETEGTFYIESVELIVDVE 1157

RESULT 8

AAR32354
 ID AAR32354 standard; protein; 1149 AA.

XX AAR32354;

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 14-JUN-1993 (first entry)

XX Coleoptera toxin from B.t. serovar japonensis variety Buibui.

XX Delta-endotoxin; larvae; Anomala cuprea; spore; crystal; Bacillus;

KW thuringiensis; virus; insect; FERM BP-3465.

XX Bacillus thuringiensis; serovar japonensis variety Buibui.

XX WO9303154-A1.

PD 18-FEB-1993.

XX 31-JUL-1992; 92WO-US006404.

XX 02-AUG-1991; 91JP-00193810.

PR 23-JUL-1992; 92US-00915203.

XX (MYCO) MYCOGEN CORP.

PA (KUBI) KUBOTA CORP.

XX Ohba M, Iwahana H, Sato R, Suzuki N, Ogiwara K, Sakanaka K;

PI Hori H, Asano S, Kawasaki T;

XX WPI; 1993-076511/09.

DR N-PSDB; AAQ36866.

XX New strain of Bacillus thuringiensis serovar japonensis - producing toxin

PT active against coleoptera larvae.

XX Claim 5; Page 28; 48pp; English.

XX The protein sequence is that of a toxin active against Coleoptera that is
 CC produced from a pure culture of Bacillus thuringiensis serovar japonensis
 CC variety Buibui (FERM BP-3465). The toxin is a delta-endotoxin which has a
 CC mol. wt. of ca. 130 kD. It is useful for control of coleoptera larvae
 CC e.g. it is effective against Anomala cuprea but has little effect on
 CC Lepidoptera. The toxin can be used as B.t. spores or crystals, as opt.
 CC treated cells (B.t. or transformed microorganisms) or it is expressed by
 CC plants. (Updated on 25-MAR-2003 to correct FN field.) (Updated on 24-OCT-
 CC 2003 to standardise OS field)

XX Sequence 1149 AA;

Query Match 60.6%; Score 3662.5; DB 2; Length 1149;

Best Local Similarity 63.8%; Pred. No. 1.1e-232;

Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

Qy 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTTLQNNMYKDYLRMSGENPELFGNP 60

Db 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTTLQNNMYKDYLRMSGENPELFGNP 60
 Qy 61 BTFTSS-STVOTGIGVQVGCALGVPAGOPASFPVPIVQGLWPSSTSVVMEIMKQVE 119
 Db 61 GTFISAQAVGTGIDIVSTIISGIGIPVLGEVFSILGSLIGLLMPSNNENNVQIPFNRVE 120
 Qy 120 DLIDQKITDSVRKTALAGLQGLDGLDVYKSLKMWLENRNDTRARSVVVTVYIALELDF 179
 Db 121 ELIDQKILDSVRSRAIADLANSRIAVEVYQNALEDWRKNPHSTRSAALVKERFCGAEAIL 180
 Qy 180 VAKIPSPAI SQEVPILLSVYAAQANHLHLLLDASIFGAENGFTPECEISTFVDQVTRTA 239
 Db 181 RTNMGSPSQTYETPFLPTFYAQAASLHLLVNRDVQIYCKENGYPQNDLDFYKEQVSYTA 240
 Qy 240 QYSDYCVKWNVTGLDKLGTNAASWLKYHFRREMTLLVLDLVALPPNVDTRYPIETTA 299
 Db 241 RYSDHCQVWVAGLNLGKGTAKQWVDYNNRFRRENNVWLDLVALPPNVDARIYPLETNA 300
 Qy 300 QLTREVYTDPIVFNRETSGGFCRRWSLN-----SDI-----SPSEVESAVIRSPHLFOIL 349
 Db 301 ELTREIPTDPV-----GSVVTGQSTLISWYDMIPALPSPSTLEN-LLRKDPDFTLL 352
 Qy 350 SEIEPFTTRAGLPLNNT-EYLEYVGVGHSIKYKNTWASSALERNYGTITSNKIKYYDLANK 408
 Db 353 QEIRMYTS---FRONGTIEYNYMGQRLTSLYIYGSSP--NKYSGVLAGEAEDIIIPVGQN 407
 Qy 409 DIFQVRSIGADLANVYAOVGVVPVASFLLDKNTGSGVGGFTYKSPHPTMVCQNTVYT 468
 Db 408 DIYRV--VWTIGRYTNLLGVNVPVP-YFNSNTQK-----TYSKP-----KQFAGGIKT 454
 Qy 469 ID---EIPPENEPLSRGYSRHLSHITSYSPKSNASSPARYGNLVPAMVTHRSADVTVTY 525
 Db 455 IDSGEELTYEN---YQSYSHVSVITSPEIKSTGTV--LGWVPFGWTHSSASRNNPIY 509
 Qy 526 SDKJTIQPVVAHTLVSGTTVIKGPQ-FTGGNILKRTSSGPLAY---TSVSVKSPLSQRY 581
 Db 510 ATKISQIPINKASRTSGGAVWNPQSGLYNGGVPWKLSGSGSQVINLRVATDAGK-ASQRY 568
 Qy 582 RARIRYASTTNLRLPVTIS-----GTRIYSINVKNMTWKGDDTLNFTDLATIGTA-F 633
 Db 569 RIRIRYASDRAGKP--TISSRSPENPATYSIAIATNTWSTNASLITYSTFAESGPINL 626
 Qy 634 TFSNYSDSLTVGADSPASGGBVYVDKFLIPVNAITFEAEEDLDVAKKAVNGLFTSKDAL 693
 Db 627 GISGSRFTDISITKEAGANLYIDRIEPIVNTLFEAEEDLDVAKKAVNGLFTNEKDAL 686
 Qy 694 QTSVTDYQVNAANLVECLSDDELIPNEKRMWDVAKKRLVQARNLLQDTGFNRINGEN 753
 Db 687 QTSVTDYQVNAANLIECLSDDELIPNEKRMWDVAKKRLVQARNLLQDTGFNRINGEN 746
 Qy 754 GWTSGTGIEVARGDVLPKDRSLRLTSAREIDTETPTLYQOIDSLLKPYTRYKLGPI 813
 Db 747 GWTSGTGIEVVEGDLVFKDRSLRLTSAREIDTETPTLYQOIDSLLKPYTRYKLGPI 806
 Qy 814 GSSQDLLEKLRHRANQIVKRVNPNLLPDVLVPSNCGGIDRCSEQQYVDANLALNNGEN 873
 Db 807 GSSQDLLEKLRHRANQIVKRVNPNLLPDVPSNCGGIDRCSEQQYVDANLALNNGEN 866
 Qy 874 GNMSSDSHAPSFHIDTGIDLNENTGIWVVKIPITNGYATLGNLELVEEGPLSGETLER 933
 Db 867 GNMSSDSHAPSFHIDTGIDLNENTGIWVVKIPITNGYATLGNLELVEEGPLSGETLER 926
 Qy 934 AQOQEQWQDQWARKRGASEKAYYAAKQAIIDRLFADYQDQKLNKSGVENSMDLAAQLVQS 993
 Db 927 AQOQEQWQDQWARKRGASEKAYYAAKQAIIDRLFADYQDQKLNKSGVENSMDLAAQLVQS 986
 Qy 994 IPYVYNDALPEIPGMNYSFTTELNRLOQANWLYDLRNLAI PNGDFRNLSDNNATSDVNV 1053
 Db 987 IPYVYNDALPEIPGMNYSFTTELNRLOQANWLYDLRNLAI PNGDFRNLSDNNATSDVNV 1046
 Qy 1054 QQLSDTSVLVIPPNNNSQVSQFTVQPNRYVLVLTARKEGCGVGYIIRDGANQTETLT 1113

Dd		1047	QQLSDTSVLVPMWNSQVSQQFTVPQPNRYRVLRTAKRGVGGVYIIRGANGQTETLTFF	1106
Qy		1114	NICDDDTGVLASDQTSYITKTVFPTSTEQWIDMSETE	1152
Dd		1107	NICDDDTGVLSTDTQTSYITKTVFPTSTEQWIDMSETE	1145
 RESULT 9 AAR51692 standard; protein; 1149 AA.				
XX	AAR51692;			
AC				
XX				
XX	16-OCT-2003 (revised)			
DT	04-NOV-1994 (first entry)			
XX				
DE	B.thuringiensis serovar Japonensis insecticidal protein.			
XX				
KW	insecticidal protein; Coleoptera larvae; Buibui fungus.			
XX				
OS	Bacillus thuringiensis; (serovar Japonensis.			
OS	strain Buibui).			
XX				
XX	Key Location/Qualifiers			
FT	Region	751..766		
FT	/note= "amino acids 751-766 are not given in the			
FT	specification and so have been decoded from AAO58975"			
XX				
PN	JPO6065292-A.			
XX				
PD	08-MAR-1994.			
XX				
PP	11-AUG-1992; 92JP-00213886.			
XX				
PR	11-AUG-1992; 92JP-00213886.			
XX	(KUBI) KUBOTA CORP.			
PA				
XX	WPI; 1994-121220/15.			
DR	N-PSDB; AAO58975.			
DR				
XX	Insecticidal protein and DNA from Bacillus thuringiensis serovar			
PT	Japonensis strain Buibui - useful in insecticides against Coleoptera			
PT	insects.			
XX				
PS	Claim 1; Page 9-13; 18pp; Japanese.			
XX				
CC	This insecticidal protein has activity against Coleopteran insect larvae			
CC	and has been isolated from Bacillus thuringiensis serovar japonensis			
CC	strain Buibui. (Updated on 16-OCT-2003 to standardise OS field)			
XX				
SQ	Sequence 1149 AA;			
 Query Match 60.6%; Score 3662.5; DB 2; Length 1149; Best Local Similarity 63.8%; Pred. No. 1.1e-232; Matches 75; Conservative 109; Mismatches 257; Indels 61; Gaps 21;				
Qy	1 MSPNNQNEYRILDASSSTSVDNSVRYPLANDQTTLQNMYKYDYLMSEGENPELEGNP	60		
Dd	: : : : : : : : : :			
Dd	1 MSPNNQNEYRIIDALSPSTSVDNSIRYPLANDQTNTLQNNNYKYDYLMTESTNAELSRNP	60		
Qy	61 ETPISS-STVOTGIIGVOVLGALGVDPFAGOLASFYSFIVGOLWPSTSVSWEMIMKOVE	119		
Dd	: : : : : : : : : :			
Dd	61 GTTISADAVGTGDIDVTSTIGLGIPVLGEVFSLGLGLWPSNENWQIFPNRVE	120		
Qy	120 DLIDQKITDSVRKTALAGLGDLGVQKSKNLMLNRNDTRARSVVVTVQVIALELDF	179		
Dd	: : : : : : : : : : :			
Dd	121 ELIDQKILDSVRSPAIADLANSRIAIEYYQNALEDKRNPHSTRSAALVKERFCNAEAIL	180		
Qy	180 VAKIPSPAISGOEVPLLSVTAQAANLHLLLRDASIFGAEWGTFPGBISTFYDQRVTRTA	239		
Dd	: : : : : : : : : : :			
Dd	181 RTNMGSFSQNYETPLLPTVAQAASHLLVWRDVQIYKKEWGPQNDIDLIFYEQVSYTA	240		

Qy	240	QYSDYCVKYNVTGLDKLKTNAASWIKYHQPREMTLLVLDLVALPNYDTRVPYIETTA	239
Db	241	RYSDHCQVMTNAGLKLGTGAKQMDYRFRREMMVWMLDLVALPNYDARLYPLETNA	300
Qy	300	QLTRREVTPDIPVFNRETSGFCRRWSLN---SDI-----SPSEVESAVIRSPHLPDIL	349
Db	301	ELTRBIFTDIV-----GSYVYQGSGSTLISWYDMIPAAALPSESTLEN-LLRXPDPFTLL	352
Qy	350	SEIEFYTTTRAGLPLANN--EYLEYVWGHISIKYKNTNASSALERNYGITTSNKIKYYDLANK	408
Db	353	QBRIMYTS---PRONGTIEYNTYMGGORULTSYIYGSSP--NKYSGVLAGEADIIPQVGN	407
Qy	409	DIPOVRSGLADLANYYAQVGVPIYASFTLLDRKTSVGGPFTYKSPHTTQWCTQNYMT	468
Db	408	DIYRV--VMYIIGRYTNSLLGVNPPVTF-YFSNNYTK-----TYSXP-----KQFAGGIKT	454
Qy	469	ID---BIPPENBPLSGYSHRLSHITSYSPSKNASSPARYGNLPVANTHRSDADVNTVY	525
Db	455	IDSGELTYEN---YQSYSHRVSYIITSPEIKSTGGTV--LGVVPIFGWTHSSASRNNPIY	509
Qy	526	SDKITQIPVVKARTLVSGTTVIKPGP-PTGNNILKRTSSGPLAY---TSVSVKSPLSQRY	581
Db	510	ATKISQIPINKASRTSGGAVNMPQEBLYNGPVMKLSGSGSOVINLRVATDAKG-ASQRY	568
Qy	582	RARIRVASTNNLRLFTIS-----GTRIYSINWKNYKNGKDDLLFPNFPDLATIGTA-F	633
Db	569	RIRIRVASDRAGKF--TISRSPPENPATYSASIAYTNTWSTNASTLYSTPAYAESGPINL	626
Qy	634	TPGNSYSDSLTVGADSPASGGEVVVDKFPFELIPVNATFEAEEDLDVAKAVNGLFTSKKDAL	693
Db	627	GISGSRTFDISITKAGAGANLYIDIRIEIPVNTLFEAEEDLDVAKAVNGLFTNEKDAL	686
Qy	694	QTSVTDYQVNAANLYECLSDSELYPNEXRMLMDAVKEAKELVQARNLLQDTGPNRRINGEN	753
Db	687	QTSVTDYQVNAANLYECLSDSELYPNEXRMLMDAVKEAKELVQARNLLQDTGPNRRINGEN	746
Qy	754	GWTSGTGIEVAGDVLFPKORSRLKLTSAREIDTETYPYLYQQIDESLLKPYTRYKLGFI	813
Db	747	GWTSGTGIEVAGDVLFPKORSRLKLTSAREIDTETYPYLYQQIDESLLKPYTRYKLGFI	806
Qy	814	GSSQDLLEIKLIRHRANOIKVNVPDNLLPDVLVNSCGGIDRCSEQQVVDANLALENNGEN	873
Db	807	GSSQDLLEIKLIRHRANOIKVNVPDNLLPDVPRVNSCGGVDRCSQQVVDANLALENNGEN	866
Qy	874	GNMSSDSHAFSPHIDTGEIDLNENTGIVWVFKIPTTTNGYATTGLENLVBEGPLSGETLER	933
Db	867	GNMSSDSHAFSPHIDTGEIDLNENTGIVWVFKIPTTTNGYATTGLENLVBEGPLSGETLEW	926
Qy	934	AQOQEQWQDMKARKGASEKATYAAKQAIIDRLFADYQDQKLSNGVEMSDMLAAQNLVQS	993
Db	927	AQOQEQWQDMKARKGASEKTYAAKQAIIDRLFADYQDQKLSNGVEMSDMLAAQNLVQS	986
Qy	994	IPVYVNDALPEIPGMNYSFTTELNLQQAOWNLYDLRNAIPNGDFPRLGSDNNATSDVNV	1053
Db	987	IPVYVNDALPEIPGMNYSFTTELNLQQAOWNLYDLQNALPNGDFPRLGSLNNNATSDVNV	1046
Qy	1054	QLSDTSVLVIPNNNSQVSCQFTVQPNRYRVLVYRTARKEGVGDGYVIIRDGANQTETLTF	1113
Db	1047	QLSDTSVLVIPNNNSQVSCQFTVQPNRYRVLVYRTARKEGVGDGYVIIRDGANQTETLTF	1106
Qy	1114	NICDDDTGVLSDADQTSYITKTVEFTPTSTQVWIDMSETE	1152
Db	1107	NICDDDTGVLSTDTQTSYITKTVEFTPTSTQVWIDMSETE	1145
RESULT 10			
AAU99256			
ID AAU99256 standard; protein; 1210 AA.			
XX			
AC AAU99256;			
XX			
DT 07-OCT-2002 (first entry)			
XX			

XX	Bacillus thuringiensis.	415	VLLDIYVPGYTYIPFGMPPEVEFFPMVQLNNTRK-----TLKYNPVSK	456
OS				
XX	WO200234774-A2.	471	-----BIPPE--NEPLSRGYSRLSHITSYFSKNASSPARYGNLPFAWTHRS	517
PN				
XX	02-MAY-2002.	457	DIATSTRDSELELPETSQDPNYESYHRLCHITSIPATGNTT-----GLVPVFSWTHRS	511
PD				
XX	24-OCT-2001; 2001WO-US045468.	518	ADVTNTVYSDKITQIPVVKANTLVSGTTVTKGPGFTCGNLIKRTSS-----GPLAYTSVSU	573
PF				
XX	24-OCT-2000; 2000US-0242838P.	512	ADLNTIYSDKITQIPAVKCDNLFPVVPVVGEGHGTGDLQYNNRSTGSGVGLFLARYGL	571
PR				
XX	23-OCT-2001; 2001US-00032717.	574	KSPLSORYRARIYASTTNLRLFTVLTSGTRIYSINVNKTNKGGDLTNTFDLATIGTAF	633
PR				
XX	(DUPO) DU PONT DE NEMOURS & CO E I.	572	ALEKAGKIRVRLAYTADADIVLHND-----QMQPKTNPGEDLTSTKTFKADAITL	626
PA				
PI	Abad AR, Duck NB, Feng X, Flanagan RD, Kahn TW, Sims LE;	634	---TFSNYSDSLTVGADSPAS-GGEVYVDKFIIPVNATFEABEDLDVAKAVNGLFTSK	689
XX	WPI; 2002-519178/55.			
XX	N-PSDB; ABR87234.	627	NLATDSSLALKHNLGHPNSTLSGIVVVDRIEPIPVDETYEAEQDLEAKKAVNALPTNT	686
DR				
XX	New isolated pesticidal polypeptide useful for impacting insect pest e.g.	690	KDALQTSVTDYQVNOAANLVECLSDLELYPNEKMLMDAVKAKELVQARNLLQDTGNRI	749
PT	Colorado potato beetle.	687	KDGLRPGVTDYEVNOAANLVECLSDLLYPNEKELLFDVAREAKELSEARNLLQDPFOI	746
PT				
XX	Claim 4; Page 96-98; 176pp; English.	750	NGENGWTSGTGIVARGDVLFDORSRLTSAREIDTETPTLYLQOIDESELLKPVTRKL	809
PS				
XX	The present invention relates to a new pesticidal polypeptide. The	747	NGENGWTSAGTGIIVIEGDALEFGRLPGAREIDTETPTLYLQKVEEGVLPKPYTRL	806
CC	invention is useful for impacting an insect pest by applying the			
CC	molecules of the invention to the environment of the insect pest by	810	KGFISQDLLEIKLIRHRANQIVKNVDPNLLPDVLPVNSCGGIDRCSEQQVYVDANLLEN	869
CC	spraying, dusting, broadcasting, or seed coating, where the insect pest			
CC	is selected from Colorado potato beetle, western corn rootworm or	807	RGFVGSQGLEIETIRHQTNRIVKNVDPDILLPDVSPVNSDGSINRCSEQKVVNSRLEVEN	866
CC	southern corn rootworm. The invention is also useful for increasing			
CC	insect target range and for producing transgenic microorganisms and	870	NGENGMSDSHAPSHPHIDTGEIDLNTGNTGIWVVPKIPTTNGYATLGNLELVEBGLSGE	929
CC	plants that express the pesticidal polypeptide. The invention is also			
CC	useful for producing transformed plants and in transforming any organism	867	R-----SGEAHFSIPIDTGEIDYENAGIWWGFKITDPEGYATLGNLELVEBGLSGD	920
CC	to produce the pesticidal polypeptide of the invention. The present amino			
CC	acid sequence represents a Bacillus thuringiensis wild-type Cry12Ib	930	TLERAOOEOWQDKMARKEGASEKAYAAKQAITDLFPADYQDOKLNSGVMSDMLAAQN	989
CC	endotoxin protein			
XX		921	ALERLQREBOOWKIQMTRRREEDRRYMAKQAVDRLYADYQDQQLPDPVEITDLTAAQD	980
XX	Sequence 1206 AA;			
SQ				
	Query Match 53.7%; Score 3246.5; DB 5; Length 1206;			
	Best Local Similarity 53.4%; Pred. No. 3.4e-205;			
	Matches 665; Conservative 172; Mismatches 290; Indels 119; Gaps 18;			
QY	1 MSPNNQNEYEILDASSTVSNSRYPLANDQTTTLQNNYKDYLMSEGENPELFGNP	60		
DB	1 MSPNNQNEYEILDATPSTVSNSRYPFANEPTNALQNDYKDYLMKSAGNASEYPCSP	60		
QY	61 ETPIIS-SSTVQTGIGVQVGLGALGVPPFAGQIASFYFVQGLWPSSTVSWEIMKQVE	119		
DB	61 EVLVSGQDAKAAIDIVGKLLSLGLGVFPVGPVIVSLYTLQDILWPSGKESQWEIFMQVE	120		
QY	120 DLIDOKITDSVRKTALAGLQGLDGVQYKSLKNWLENDRTRARSVVVTVYALELDF	179		
DB	121 ELINQKIAEYARKAUSLEBGLGNVYQYLTALBEEENPENGSKALDRVRREILDSLP	180		
QY	180 VAKIPSPAISQEVPLLSVYAQANLHLLLRDASIFCAEWGFTPKGEISTFYDRQVTRTA	239		
DB	181 TQMPSPRVNFPVFFITVYMANLHLLLKQASIFGEESGWSSTTINNYDRQMKLTA	240		
QY	240 QYSDYCVKMYNTGLDKLKGNTAASWLKHYFRREMTLLVLDLVALPNYDTRTYPIBTTA	299		
DB	241 EYSDHCVKMYETGLAKLKGTSKQWVDYQFRREMTLAVLDVVALPNYDTRTYPIBTTA	300		
QY	300 QLREVVYTDPIVFNRETSFGCFRWSLNSDISSEVESAVIRSPHLFDILSEIBFYTTA	359		
DB	301 QLREVVYTDPIVFNRETSFGCFRWSLNSDISSEVESAVIRSPHLFDILSEIBFYTTA	356		
QY	360 GLPLNTEYLEYVWGHSHIKYKNTNASSALERNYGTITS-NKIKYDILANKDIFQVRSIGA	418		
DB	357 S--ISSARYIRHWAGHQSIVHVSRSGLNQMYCTGNQLHSTSTFDFTNYDIYKLSKDA	414		
QY	419 DLAN-----YYAQVGVYPYASFTLLDK--NFGSGSVGGFTYKPHHTMQVCTQYNTID-	470		

RESULT 12

AAW06417

ID AAW06417 standard; protein; 1169 AA.

XX AAW06417;

AC AAW06417;

XX 16-OCT-2003 (revised)

DT 28-JAN-1997 (first entry)

XX Antiscarab pest toxin 50C(b).

XX Toxin 50C(b); scarab pest; toxin; insect; scarabaeidae; pest control;

XX larval stage insect; grain; tuberous crop; white grub; chafer grub;

XX cyclocephala; popillia.

XX Bacillus thuringiensis; strain kumamotoensis.

XX US5554534-A.

PN

PF 25-JUN-1998; 98MO-EP004033.
XX
PR 27-JUN-1997; 97US-00884389.
XX
XX (PLBZ) PLANT GENETIC SYSTEMS NV.
PA
XX De Roock S, Van Rie J;
PI
XX WPI; 1999-105666/09.
DR
XX
XX New Bacillus thuringiensis Cry9C crystal proteins with improved toxicity
PT - useful for producing transgenic plants in controlling pest insects,
PT especially Ostrinia nubilalis, Heliothis virescens and Diatraea
PT grandiosella.
XX
XX Example; Page; 40pp; English.
PS
XX This is the amino acid sequence of a Cry9C (crystal) mutant protein which
CC is toxic to an insect species. It is modified in the method of the
CC invention in an attempt to improve its toxicity. The improved proteins
CC are useful for producing transgenic plants and seeds in controlling
CC insects feeding on a plant. The transgenic plants are also useful for
CC producing more transgenic plants with similar characteristics, or for
CC introducing the improved cry9C gene (including fragments) into the same
CC plant variety or related plant species. The Cry9C protein can be obtained
CC from cultured cells of the transgenic plants and used in insecticide
CC compositions against lepidopteran insects, especially Ostrinia nubilalis,
CC Heliothis virescens and Diatraea grandiosella. The cry9C gene can be
CC expressed in combination with another insect control protein (e.g.
CC Another B.thuringiensis derived crystal protein, especially CryIb- or
CC CryIIb-type protein), and used to transform B.thuringiensis, which
CC produces other insecticidal toxins. Such strains are useful against a
CC variety of insect pests, or insects where insect resistance development
CC is prevented or delayed. (Note: this sequence is not given in the
CC specification, but is generated using the information provided by the
CC inventors.)
XX
XX SQ Sequence 1157 AA;

Query Match 51.7%; Score 3126; DB 2; Length 1157;
Best Local Similarity 52.3%; Pred. No. 2.9e-197;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

QY 1 MSPNNQNEYIELDASSTSVDSNRYPLANDQTTTLQNNYKYDYLMSRGENPELPGNP 60
DB 1 MNENNQNEYEIIIDAPHCFCFSDDDVRYPLASDPNAALQNNYKYDYLQMTDEYDTSYINP 60
QY 61 ETPISS-STVQTGIGIVGVGLGALGVPFAGQIASFYSFIVGQLMPSTSVSWEMIKQVE 119
DB 61 SLISGRDAVQTALTUVVGRITGALGVFPSCQIVSFYQFLTLTPVNDTAIMEAFMRQVE 120
QY 120 DLIDDKITDSVRKTALAGLGLDGLDYQKSLKNLKNLNDTRARSVVVTVQVIALELDF 179
DB 121 ELVNNQITEFARQALARLQGLGDSFNRYQSLQNLADNRDNLNSVRAQFIALDLDF 180
QY 180 VAKIPSPAISQSQVPLLSVYAQANLHLLLRDASIFGAEMGFTPGHISTFYDRQVTRTA 239
DB 181 VNAIPLFAVNGQVPLLSVYAQANLHLLKDDASLFGEGMGFTQGBISTFYDRQLelta 240
QY 240 QYSDYCVKRWNTGLDKGTNAASLWKHOFREMTLLVLDLVALFPNYTRTPPIETTA 299
DB 241 KYTNYCETWTNTGLDRLGRNTESWLRHOFREMTLVLDLVALFPNYDRLYPTGNSP 300
QY 300 QLTREVVYTDPIVNRRETSQGFRCRWNSINDISFSEVASVIRSPHLPDILSBEIEFTTRA 359
DB 301 QLTREVVYTDPIVFNPPNAGLCRWNGTNPNTYSELENAFIRPHLPDLRLNSITISNR - 359
QY 360 GLPLNNTEYLEYVGHSHIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFQVRSI 416
DB 360 -FPV-SNFPMDYWSGHTLRSYLDNSAVQEDSYGLITTTTRATINPGVQGTNR---IEST 413
QY 417 GADLANYAQQVGVPIYASFTL--LDKNVSGSGVGGFTYKPHMTMQVCTQNYNTIDIEPP 474

Db 414 AVDFRSALIGYGNRASFPVPGGLFNCTTSPANGG-----CRDLYDNDLPP 461
QY 475 ENEPLSRGYS-HRLSHITSYSFSKN-ASSPARYNLVPFAWTHRSADVTNTVYSDKITQI 532
Db 462 DE---STGSSTHRLSHVTFPSFQTNQAGSIANAGSVTVYVTRRDVLDLNTITPNIITOL 518
QY 533 PVYKAHTLVSGTTVIKGPGFTGNNILKRTSGPLAYTSVSKSPLSQRYARIRYASTTN 592
Db 519 PLVKASAPVSGTTLVKPGFTGGIILRTTNGTFTGLRVTVNSELTOQYRLRVRPASTGN 578
QY 593 LRLPVITSGPRIYSINWNKTMKGGDLTENTF---DLATIG---TAPFESNYSLSLTGCA 646
Db 579 PSIRVLRGVSGISGDVRLGSTMNRGQLETSFPTRFTTTGPPPPFTTQAOQLTVA 638
QY 647 DSPASGEVYVVKFELIPVNATFEABEDLDVAKAVNGLPFTSKKDALQTSVTDYQVNOAA 706
Db 639 EGVSTGGEYIIDRIEIVPNPAREABEDLEAAKAVASLPTRTDRGLQVNVTDYQVQAA 698
QY 707 NLVTECLSDLEYPNEKMLMDAVKEAKELVQARNLQDTGPNRING--ENGWGTSGTIEVA 764
Db 699 NLVSCLSDEQYGHDKMLLEAVRAAKRLSERNLQDPDFTNTINSTENGWKAANGVTIS 758
QY 765 EGDVLPKORSRLTSAEIDTETVPTLYQOIDSLLKPYTRYKLGFGISSQDLLEIKLI 824
Db 759 EGGPFFAGRALQASAR----ENPTTIIYQKVDASVLKPYTRYRDLGFPVKSQDLLEIDL 814
QY 825 RHRANDITKQVNPONLPLPDLVFNASCGRSCQYVDANLALNENGENGNMSSDSHAPS 884
Db 815 HHKVVHLVKNVDPNLVSDTYSDCSGINRCDEQHQVDMQDAEHPMDCCAAQTHERFS 874
QY 885 FHIDTGBIDLNWTVWVVKIPTTNGYATLGNLELVEBGLSGETLERAQOQOQOQWQDK 944
Db 875 SYINTGDLNASVDQGIWVVLKVRTTDGYATLGNLELVEVGLSGESLEREQDRDNKMAE 934
QY 945 MAREKRGASEKAYAAKQAIIDRLPADYODOKLNSGVEMSDMLAQNLYQSIPIVYVNDALPE 1004
Db 935 LGRKRABIDRVYLAAKQAINHLFDYQDQQLNPEIGLABINEASNLVESISGVYSDTLIQ 994
QY 1005 IPGMYNTSPTTELNRLOQAWNLVLRNAIPNGDFRNLGSDWNATSDVNVQOQSDTSVLVI 1064
Db 995 IPGINVEIYTELSDRLQQAQSYLYTSRNAVQNGDFNSGLDSWNTTMDASVQDQGNHFLVL 1054
QY 1065 PNWNSQVSQOFTVQPNRYVLRVTARKEGVDGVVIRIDGANQOTETLTFTNI CDDDTGVLS 1124
Db 1055 SHMDAQVSOQLRVNPNCKYVLRVTARKVGGGQGVYVTRIDGAHQETLTFTNACDYVNGTY 1114
QY 1125 ADQTSYITTKVETPSTEQVWIDMSETEGVFNIESVELVLEBE 1167
Db 1115 VNDNSYITERVVPETKEMWVEVSESGSFYDSDIEFIETQE 1157

Search completed: June 21, 2004, 10:14:49

Job time : 98 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:12:40 ; Search time 402 Seconds
(without alignments)
819.552 Million cell updates/sec

Title: US-10-089-678-1

Perfect score: 6044

Sequence: 1 MSPNNQNEYILDASSSTSV.....MSETGVNIESVELVLEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6044	100.0	1167	12 US-10-089-678-1	Sequence 1, Appli
2	3263.5	54.0	1210	13 US-10-032-717-4	Sequence 4, Appli
3	3263.5	54.0	1210	14 US-10-414-637-4	Sequence 4, Appli
4	3263.5	54.0	1210	16 US-10-606-320-4	Sequence 2, Appli
5	3246.5	53.7	1206	13 US-10-032-717-2	Sequence 2, Appli
6	3246.5	53.7	1206	14 US-10-414-637-2	Sequence 2, Appli
7	3246.5	53.7	1206	16 US-10-606-320-2	Sequence 2, Appli
8	3100.5	51.3	1156	14 US-10-099-285-72	Sequence 72, Appli
9	3100.5	51.3	1156	15 US-10-428-961-28	Sequence 28, Appli
10	2884.5	47.7	1152	15 US-10-614-524-6	Sequence 6, Appli
11	2825.5	46.7	1150	14 US-10-099-285-74	Sequence 74, Appli
12	2710	44.8	1134	14 US-10-099-285-76	Sequence 76, Appli
13	2519.5	41.7	1207	10 US-09-988-462-7	Sequence 7, Appli
14	2457	40.7	1227	15 US-10-428-961-63	Sequence 63, Appli
15	2454.5	40.6	1186	9 US-09-826-660-23	Sequence 23, Appli

16	2449.5	40.5	1228	15 US-10-428-961-38	Sequence 38, Appli
17	2446.5	40.5	1228	15 US-10-614-524-2	Sequence 2, Appli
18	2230.5	36.9	1177	9 US-09-873-873-10	Sequence 10, Appli
19	2230.5	36.9	1177	9 US-09-873-873-12	Sequence 12, Appli
20	2230.5	36.9	1177	9 US-09-873-873-14	Sequence 14, Appli
21	2230.5	36.9	1177	10 US-09-997-914-10	Sequence 10, Appli
22	2230.5	36.9	1177	10 US-09-997-914-12	Sequence 12, Appli
23	2230.5	36.9	1177	10 US-09-997-914-14	Sequence 14, Appli
24	2230.5	36.9	1177	14 US-10-365-645-10	Sequence 10, Appli
25	2230.5	36.9	1177	14 US-10-365-645-12	Sequence 12, Appli
26	2230.5	36.9	1177	14 US-10-365-645-14	Sequence 14, Appli
27	2230.5	36.9	1177	16 US-10-672-163-10	Sequence 10, Appli
28	2230.5	36.9	1177	16 US-10-672-163-12	Sequence 12, Appli
29	2230.5	36.9	1177	16 US-10-672-163-14	Sequence 14, Appli
30	2229.5	36.9	1177	9 US-09-873-873-26	Sequence 26, Appli
31	2229.5	36.9	1177	10 US-09-997-914-26	Sequence 26, Appli
32	2229.5	36.9	1177	14 US-10-365-645-26	Sequence 26, Appli
33	2229.5	36.9	1177	16 US-10-672-163-26	Sequence 26, Appli
34	2227.5	36.9	1163	9 US-09-756-526A-2	Sequence 2, Appli
35	2227.5	36.9	1163	14 US-10-345-020-2	Sequence 2, Appli
36	2227.5	36.9	1163	14 US-10-342-821-2	Sequence 2, Appli
37	2225.5	36.8	1177	9 US-09-873-873-28	Sequence 28, Appli
38	2225.5	36.8	1177	10 US-09-997-914-28	Sequence 28, Appli
39	2225.5	36.8	1177	14 US-10-365-645-28	Sequence 28, Appli
40	2225.5	36.8	1177	16 US-10-672-163-28	Sequence 28, Appli
41	2215.5	36.7	1177	9 US-09-873-873-34	Sequence 34, Appli
42	2215.5	36.7	1177	14 US-10-365-645-34	Sequence 34, Appli
43	2191	36.3	1156	14 US-10-099-285-70	Sequence 70, Appli
44	2154.5	35.6	1174	9 US-09-826-660-6	Sequence 6, Appli
45	2154.5	35.6	1174	12 US-09-837-961-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-10-089-678-1
; Sequence, Application US/10089678
; Publication No. US20030017967A1
; GENERAL INFORMATION:
; APPLICANT: ASANO, Shinichiro
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, AND METHOD
; FILE REFERENCE: Q68821
; CURRENT APPLICATION NUMBER: US/10/089,678
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: JP 2000-236140
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/JP01/06660
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-089-678-1

Query Match	100.0%	Score 6044;	DB 12;	Length 1167;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1167;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSPNNQNEYILDASSSTSVNSVRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP	60	
Db	1	MSPNNQNEYILDASSSTSVNSVRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP	60	
QY	61	ETPFSSSTVOTGIGIVGVGLGALGVPPFAGQIASFSYPIVQQLMPSSTVSWEMIMQVED	120	
Db	61	ETPFSSSTVOTGIGIVGVGLGALGVPPFAGQIASFSYPIVQQLMPSSTVSWEMIMQVED	120	
QY	121	LIDQKITTSVRKLTALAGLQGLDGVYQKSLKNWLENENDTRARSVVVTVYIALDLDFV	180	
Db	121	LIDQKITTSVRKLTALAGLQGLDGVYQKSLKNWLENENDTRARSVVVTVYIALDLDFV	180	


```
QY 181 AKIPSAISQGVPLLSVVAQAANLHLLLRDASIFGAEWGFTPGBISTFYDQVTRTAQ 240
DB 181 AKIPSAISQGVPLLSVVAQAANLHLLLRDASIFGAEWGFTPGBISTFYDQVTRTAQ 240
QY 241 YSDYCVKWTNTGLDKLKGNTAAASWLKHYQFRREMTLLVLDLVALFPNYDTRTYPIETTTAQ 300
DB 241 YSDYCVKWTNTGLDKLKGNTAAASWLKHYQFRREMTLLVLDLVALFPNYDTRTYPIETTTAQ 300
QY 301 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISSEVESAVIRSPHLPDILSEIEFYTRAG 360
DB 301 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISSEVESAVIRSPHLPDILSEIEFYTRAG 360
QY 361 LPLANNTEYLEYVGHSHIKYKNTNASSALERNYGTITSNKIKYDYLANKDIFQVRSIGADL 420
DB 361 LPLANNTEYLEYVGHSHIKYKNTNASSALERNYGTITSNKIKYDYLANKDIFQVRSIGADL 420
QY 421 ANYAQVYGVPIASFTLLDKNTGSGVGGFTYSKPHHTMQVCTQNTYNTIDEIPPENEPLS 480
DB 421 ANYAQVYGVPIASFTLLDKNTGSGVGGFTYSKPHHTMQVCTQNTYNTIDEIPPENEPLS 480
QY 481 RGVSHRLSHITSYSPSKNASSPARYGNLPPVATWHSADVTNTVYSDKITQIPVVKARHL 540
DB 481 RGVSHRLSHITSYSPSKNASSPARYGNLPPVATWHSADVTNTVYSDKITQIPVVKARHL 540
QY 541 VSGTIVIKGPGFTGGNLIKRTSSGPLAYTSVSVKSPLSQRYRARIYASTTNLRLFWTIS 600
DB 541 VSGTIVIKGPGFTGGNLIKRTSSGPLAYTSVSVKSPLSQRYRARIYASTTNLRLFWTIS 600
QY 601 GTRIYSINVNKTNKGGDDLTFTNTPDLATIGTAPTSNYSDSLTVGADSPASGEVYVDFP 660
DB 601 GTRIYSINVNKTNKGGDDLTFTNTPDLATIGTAPTSNYSDSLTVGADSPASGEVYVDFP 660
QY 661 ELIPNATPEABEDLDVAKAVNGLFTSKDALQTSVTDYQVQANLVKESDSELYPNE 720
DB 661 ELIPNATPEABEDLDVAKAVNGLFTSKDALQTSVTDYQVQANLVKESDSELYPNE 720
QY 721 KRLMDVAKEAKRLQVARNLLQDTGFNRINGENGTGSGTIEVABGDVLFKDRSLRLTSA 780
DB 721 KRLMDVAKEAKRLQVARNLLQDTGFNRINGENGTGSGTIEVABGDVLFKDRSLRLTSA 780
QY 781 REIDTETPTLYLQYIDESLLKPYTRYKLKGFIGSSQDLEIKLIRHRANQIVKQVDPNLL 840
DB 781 REIDTETPTLYLQYIDESLLKPYTRYKLKGFIGSSQDLEIKLIRHRANQIVKQVDPNLL 840
QY 841 PDVLPVNSCGGIDRCSEQOYVDANLALENNGENGNMSSDHAFTSHIDTGEIDLNENTGI 900
DB 841 PDVLPVNSCGGIDRCSEQOYVDANLALENNGENGNMSSDHAFTSHIDTGEIDLNENTGI 900
QY 901 WYVFKIPTTNGVATLGNLEVBEGPLSGETLERAQQEQOQWQDKMARKEGASEKAYYAAK 960
DB 901 WYVFKIPTTNGVATLGNLEVBEGPLSGETLERAQQEQOQWQDKMARKEGASEKAYYAAK 960
QY 961 QAIIDLFPADYQOKLNSGVENSDMLAAQNLVQSIPIVYNDALPEIPGMNYTTFTELTRNL 1020
DB 961 QAIIDLFPADYQOKLNSGVENSDMLAAQNLVQSIPIVYNDALPEIPGMNYTTFTELTRNL 1020
QY 1021 QQANWLYDLRNAIPNGDFRGLSDNNATSDVNVQQLSDTSVLVIPNWSQVSOQFTVQBN 1080
DB 1021 QQANWLYDLRNAIPNGDFRGLSDNNATSDVNVQQLSDTSVLVIPNWSQVSOQFTVQBN 1080
QY 1081 YRYVLRTARKEGVGVDGYVIRIDGANQOTETLAFNICDDDTGVLSDAQTSYITKTVEFTPS 1140
DB 1081 YRYVLRTARKEGVGVDGYVIRIDGANQOTETLAFNICDDDTGVLSDAQTSYITKTVEFTPS 1140
QY 1141 TEQWIDMSSETGVFNIESVELVLEE 1167
DB 1141 TEQWIDMSSETGVFNIESVELVLEE 1167
```

RESULT 2
US-10-032-717-4
; Sequence 4, Application US/10032717

```
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins With
; Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/032,717
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; TYPE: PRN
; ORGANISM: Bacillus thuringiensis
; US-10-032-717-4

Query Match 54.0%; Score 3263.5; DB 13; Length 1210;
Best Local Similarity 53.4%; Pred. No. 3e-259;
Matches 657; Conservative 178; Mismatches 284; Indels 119; Gaps 21;

QY 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMGEGENPELFGNP 60
DB 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMGEGENPELFGNP 60
QY 61 EFTFIS-SSVQTCIGIGVQVLAGALGVFPAGQIASFYFVQGLWPSSTSVSWEMIMKQVE 119
DB 61 EFTFIS-SSVQTCIGIGVQVLAGALGVFPAGQIASFYFVQGLWPSSTSVSWEMIMKQVE 119
QY 120 DLIDDKITDSVKTALAGLQGLDGLDVYQKSLQWLENRNDTRASVVTQVIAELDP 179
DB 120 DLIDDKITDSVKTALAGLQGLDGLDVYQKSLQWLENRNDTRASVVTQVIAELDP 179
QY 180 VAKIPSPALSGOEVPLLSVYQAANLHLLLRDASIFGAEWGFTPGBISTFYDQVTRTA 239
DB 180 VAKIPSPALSGOEVPLLSVYQAANLHLLLRDASIFGAEWGFTPGBISTFYDQVTRTA 239
QY 240 QYSDYCVKWTNTGLDKLKGNTAAASWLKHYQFRREMTLLVLDLVALFPNYDTRTYPIETTTAQ 299
DB 240 QYSDYCVKWTNTGLDKLKGNTAAASWLKHYQFRREMTLLVLDLVALFPNYDTRTYPIETTTAQ 299
QY 300 QLTREVYVTDPIVFNRETSGGFCRRWSLNSDISSEVESAVIRSPHLPDILSEIEFYTRTA 359
DB 300 QLTREVYVTDPIVFNRETSGGFCRRWSLNSDISSEVESAVIRSPHLPDILSEIEFYTRTA 359
QY 360 GLPLANNTEYLEYVGHSHIKYKNTNASSALERNYGTITSNKIKYDYLANKDIFQVRSIGLA 418
DB 360 GLPLANNTEYLEYVGHSHIKYKNTNASSALERNYGTITSNKIKYDYLANKDIFQVRSIGLA 418
QY 419 DIAN-----YAAQVYGVPIASFTLLDK--NTSGSGVGGFTYSKPHHTMQVCTQNTYNTID- 470
DB 419 DIAN-----YAAQVYGVPIASFTLLDK--NTSGSGVGGFTYSKPHHTMQVCTQNTYNTID- 470
QY 471 VLLDIVPFGYTYTIFFGMEVEBPFMVNQLANTRK-----TLKYNPVS 456
DB 471 VLLDIVPFGYTYTIFFGMEVEBPFMVNQLANTRK-----TLKYNPVS 456
QY 497 -----PIPPE--NEPLSRGVSHRLSHITSYSPSKNASSPARYGNLPPVATWHS 517
DB 497 -----PIPPE--NEPLSRGVSHRLSHITSYSPSKNASSPARYGNLPPVATWHS 517
QY 518 ADVNTVYSDKITQIPVVKARHL---VSG---TTVIKPGFTGCGNLTSSGLAYTSV 571
DB 518 ADVNTVYSDKITQIPVVKARHL---VSG---TTVIKPGFTGCGNLTSSGLAYTSV 571
QY 572 SVK-SPLSQRYRARIYASTTNLRLFWTISGRTRIYSINVNKTNKGGDDLTFTNTPDLATIG 630
DB 572 SVK-SPLSQRYRARIYASTTNLRLFWTISGRTRIYSINVNKTNKGGDDLTFTNTPDLATIG 630
QY 631 TAFTTSNYSDSLTVGADSPAS---GGEVYVDFKELIPVATWHSADVTNTVYSDKITQIPVVKARHL 687
DB 631 TAFTTSNYSDSLTVGADSPAS---GGEVYVDFKELIPVATWHSADVTNTVYSDKITQIPVVKARHL 687
```


Qy	1108	TEILTFNFCDDDT-GVLUSA-----	1125
		: :	
Db	1103	TETLTFSADYDNGMYNTQVSTNGTNTNAVNTQASTNGYNANNMNTQASNTNGYN	1162
		: :	
Qy	1126	-----DOTSVITKTVEPTSTEQWIDMSCTGCVNIESTVELVLEEB	1167
		: :	
Db	1163	TNSVNDQTGYTKVTPIPYTDQMWIEMSETGTFFYIESVELIVDVE	1210
		: :	
 RESULT 4			
US-10-606-320-4			
; Sequence 4, Application US/10606320			
; Publication No. US20040091505A1			
; GENERAL INFORMATION:			
; APPLICANT: Andre R. Abad			
; APPLICANT: Ronald D. Flannagan			
; APPLICANT: Rafael Herrmann			
; APPLICANT: Theodore W. Kahn			
; APPLICANT: Albert L. Lu			
; APPLICANT: Billy Fred McCutchen			
; APPLICANT: James K. Presnail			
; APPLICANT: James F.H. Wong			
; APPLICANT: Cao-Guo Yu			
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal			
; FILE OF INVENTION: Activity			
; FILE REFERENCE: 35718/263948			
; CURRENT APPLICATION NUMBER: US/10/606,320			
; CURRENT FILING DATE: 2003-06-25			
; PRIOR APPLICATION NUMBER: 60/391,786			
; PRIOR FILING DATE: 2002-06-26			
; PRIOR APPLICATION NUMBER: 60/460,787			
; PRIOR FILING DATE: 2003-04-04			
; NUMBER OF SEQ ID NOS: 134			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 4			
; LENGTH: 1210			
; TYPE: PRT			
; ORGANISM: Bacillus thuringiensis			
US-10-606-320-4			
Query Match 54.0%; Score 3263.5; DB 16; Length 1210;			
Best Local Similarity 53.4%; Pred. No. 3e-258;			
Matches 667; Conservative 178; Mismatches 284; Indels 119; Gaps 21;			
Qy	1	MSPNNNEYELDPASSSTSVDNSVRYPLANDOTTTLONNYKDYLRMSGENPELFGNP	60
		: :	
Db	1	MSPNNNEYELIIPATPSTVSNDNSRYPFANBPINALQNMDYKDYLRMSAGNASSEPGSP	60
		: :	
Qy	61	ETPIS--SSTVQTGIGVGVLGALGVPPFAQIASFYSGQLWPSSSTVSVMEMIKQVE	119
		: :	
Db	61	EVLVSGODAKAAIDIVGLKSLGLVPFGVPSLYTLQILDLPWGSKSQWEIFMEQVE	120
		: :	
Qy	120	DLIDQKITSVRKTAGLGGLGDLDVYQKSLKNWLNRNDTRASKVVTVYIALELDF	179
		: :	
Db	121	ELINQKIAYARNKALSELGLGNYYQLYLTALEEWKENPNGSRALDRVNRREILDSLF	180
		: :	
Qy	180	VAKIPSPALSGQEVPLISVYAQAHLHLLRDASTFGAEWGTTPGEISTFPYDQVTRTA	239
		: :	
Db	181	TQMPSPFRVTNFVPLTVTYTOANLHLLKDAISFGEEGWGSTTTINNYYDRQMKLTA	240
		: :	
Qy	240	OYSDYCKVNTTGLDKLKGTHAASWLKYHQFRREMILLVLDVALPNVDTRYPIETTA	299
		: :	
Db	241	EYSDHCVKVYENGUALKLGHSAKWQVDYNGQFRREMILTLDVVALPNVDTRYPMETKA	300
		: :	
Qy	300	QLTREVVTDPIVFNRRETSGGFCRRWSLNLSDISFSEVESAVIRSFPHLPDILSEIEPTYTRA	359
		: :	
Db	301	QLTREVVTDPLGAVNVSSIG---SW-YDKAPSGVIESSVIRPPHPVDITYGLTVYTQSR	356
		: :	
Qy	360	GLPLNNTYLEYVVGHSIKYKTNNTVAASALERNTGTTTS-NKI KYDLANKDIPQVRSIGA	418
		: :	
Db	357	S--ISSARYTRHWAGHOISYHRIFSNIIKOMYGTQNHLHSTSTFDFTNYDIYKTL SKDA	414
		: :	

```
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-10-032-717-2

Query Match      53.7%; Score 3246.5; DB 13; Length 1206;
Best Local Similarity 53.4%; Pred No. 7.5e-257;
Matches 665; Conservative 172; Mismatches 290; Indels 119; Gaps 18;

Qy 1 MSPNNQNEYIILDASSSTVSNSVRYPLANDOTTTLQNNYKDYLRMSGEPNLPFGNP 60
Db 1 MSPNNQNEYIILDASSSTVSNSVRYPLANDOTTTLQNNYKDYLRMSGEPNLPFGNP 60
Qy 61 EFTFIS-SSTVOTGIGIVGVLGALGVPPAGQIASFSYFIVGQLWPSSTVSWMIMKQVE 119
Db 61 EFTFIS-SSTVOTGIGIVGVLGALGVPPAGQIASFSYFIVGQLWPSSTVSWMIMKQVE 119
Qy 61 EVLVSGQDAKAAIDIVGKLLGLGVFPVGPVIVSLYTLQIDILWPSGKSEQMBIFMEQVE 120
Db 61 EVLVSGQDAKAAIDIVGKLLGLGVFPVGPVIVSLYTLQIDILWPSGKSEQMBIFMEQVE 120
Qy 120 DLIDOKITDSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVYIALELDF 179
Db 120 DLIDOKITDSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVYIALELDF 179
Qy 121 ELINQKIAETARNKALSELGLGNVYQLYLTALBEEENPNGSRALRDVRRFEILDSLP 180
Db 121 ELINQKIAETARNKALSELGLGNVYQLYLTALBEEENPNGSRALRDVRRFEILDSLP 180
Qy 180 VAKIPSFASISQEVPLLSVYAAQANLHLLLRDASIFGAEWGFTPGIBISTFYDQVTRTA 239
Db 180 VAKIPSFASISQEVPLLSVYAAQANLHLLLRDASIFGAEWGFTPGIBISTFYDQVTRTA 239
Qy 181 TQWPSFRVNTFVFPFLTVYMAANLHLLLRDASIFGEENGWSTTTNNYIDRQMKLTA 240
Db 181 TQWPSFRVNTFVFPFLTVYMAANLHLLLRDASIFGEENGWSTTTNNYIDRQMKLTA 240
Qy 240 QYSDYCVKNTGTLGDKLKTNAASWLKYHOFREMLLLVLDVALPNYDTRTYPIETTA 299
Db 240 QYSDYCVKNTGTLGDKLKTNAASWLKYHOFREMLLLVLDVALPNYDTRTYPIETTA 299
Qy 241 EYSDHCYKWTETGLAKLGTSAKQWYDYNQFRREMLAVLDVVALPNYDTRTYPIETRA 300
Db 241 EYSDHCYKWTETGLAKLGTSAKQWYDYNQFRREMLAVLDVVALPNYDTRTYPIETRA 300
Qy 300 QLTREVYTDPIVFNRETSRGCRRWSNSDISPSEVESAVIRSPHLDILSEIEFTYTRA 359
Db 300 QLTREVYTDPIVFNRETSRGCRRWSNSDISPSEVESAVIRSPHLDILSEIEFTYTRA 359
Qy 301 QLTREVYTDPLGANVSSIG--SW-YDKAPSGFVIESSVIRPHFDYITGLTVYTSQR 356
Db 301 QLTREVYTDPLGANVSSIG--SW-YDKAPSGFVIESSVIRPHFDYITGLTVYTSQR 356
Qy 360 GLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIYKDYLANDKDIFQVRSIGA 418
Db 360 GLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIYKDYLANDKDIFQVRSIGA 418
Qy 357 S--ISSARIYRWAGHQISYHRVSRGSLQOMYGTQNLHSTSTFDTNVDIVKTLSDKA 414
Db 357 S--ISSARIYRWAGHQISYHRVSRGSLQOMYGTQNLHSTSTFDTNVDIVKTLSDKA 414
Qy 419 DLAN-----YAAVYGVYASFTLLDK--NTGSGSGVGYTSRPHHTMQVCTQNYTID- 470
Db 419 DLAN-----YAAVYGVYASFTLLDK--NTGSGSGVGYTSRPHHTMQVCTQNYTID- 470
Qy 415 VLADIVVPGYTYIFPGAPVEFEFFWVNLNTRK-----TLKYNPVSK 456
Db 415 VLADIVVPGYTYIFPGAPVEFEFFWVNLNTRK-----TLKYNPVSK 456
Qy 471 -----ELPPE--NEPLSRGYSRHLSHITSYSFSKNASSPARYGNLPFAWTHRS 517
Db 471 -----ELPPE--NEPLSRGYSRHLSHITSYSFSKNASSPARYGNLPFAWTHRS 517
Qy 457 DIIASTRDSLELPPETSDQPNYESYSHRICHITSIPATGNTI-----GLVPFVSWTHRS 511
Db 457 DIIASTRDSLELPPETSDQPNYESYSHRICHITSIPATGNTI-----GLVPFVSWTHRS 511
Qy 518 ADVNTVYSKTIQIPVVKAAHTLVSGTIVKGRGFTGCGNLIKRTSS-----GPLAYTSVSV 573
Db 518 ADVNTVYSKTIQIPVVKAAHTLVSGTIVKGRGFTGCGNLIKRTSS-----GPLAYTSVSV 573
Qy 512 ADLNTIYSKTIQIPVVKAWDNLFPVVPVVKGRGHTGCDLLQYNRSVGTGLFLARYGL 571
Db 512 ADLNTIYSKTIQIPVVKAWDNLFPVVPVVKGRGHTGCDLLQYNRSVGTGLFLARYGL 571
Qy 574 KSPLSQRYRARIYASTNRLRFLVTSIGTRIYSINVKNTWKNGDDLTFTNFDLATIGTAP 633
Db 574 KSPLSQRYRARIYASTNRLRFLVTSIGTRIYSINVKNTWKNGDDLTFTNFDLATIGTAP 633
Qy 572 ALEKAGKRYRVLRYATDADIVLVHND-----QIQMPKTNWPGEDLTSKTFKVADAITTL 626
Db 572 ALEKAGKRYRVLRYATDADIVLVHND-----QIQMPKTNWPGEDLTSKTFKVADAITTL 626
Qy 634 ---TFPNVSDSLTVGADSPAS-CGEVVYVDKFLIPVNATFEAREDDLOVAKAVNGLFTSK 689
Db 634 ---TFPNVSDSLTVGADSPAS-CGEVVYVDKFLIPVNATFEAREDDLOVAKAVNGLFTSK 689
Qy 627 NLATDTSALAKHNLGEPDNTSLGIVYVDRIEPIFVDTEYREQDLEAAKAVNALFTNT 686
Db 627 NLATDTSALAKHNLGEPDNTSLGIVYVDRIEPIFVDTEYREQDLEAAKAVNALFTNT 686
Qy 690 KDALQTSVTVQVNOAANLVECLSDLELYPNEKRLMDAVKEAKELVOARNLLQDTGNRI 749
Db 690 KDALQTSVTVQVNOAANLVECLSDLELYPNEKRLMDAVKEAKELVOARNLLQDTGNRI 749
Qy 687 KDGLRPGVTDYEVNOAANLVECLSDLELYPNEKRLLPDAVREAKELSEARNLLQDPDQEI 746
Db 687 KDGLRPGVTDYEVNOAANLVECLSDLELYPNEKRLLPDAVREAKELSEARNLLQDPDQEI 746
Qy 750 NGENGWGTSIGIEVABGCVLPKORSRLRLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 809
Db 750 NGENGWGTSIGIEVABGCVLPKORSRLRLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 809
Qy 747 NGENGWGTSIGIEVABGCVLPKORSRLRLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 806
Db 747 NGENGWGTSIGIEVABGCVLPKORSRLRLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 806
Qy 810 KGFIGSSQDLIELKIRHANCQIVKRVNPNLLPDVLVUNSCGIDRCSEQQYVDANTALEN 869
Db 810 KGFIGSSQDLIELKIRHANCQIVKRVNPNLLPDVLVUNSCGIDRCSEQQYVDANTALEN 869
Qy 807 RGFVSSQGLEIETIRHQTRNRIKVNVPDDLPLPDVSPVNSDGSINRCSQKYVNSRLEVEN 866
Db 807 RGFVSSQGLEIETIRHQTRNRIKVNVPDDLPLPDVSPVNSDGSINRCSQKYVNSRLEVEN 866

RESULT 6
US-10-414-637-2
; Sequence 2, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn B. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414,637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032,717
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-10-414-637-2

Query Match      53.7%; Score 3246.5; DB 14; Length 1206;
Best Local Similarity 53.4%; Pred No. 7.5e-257;
Matches 665; Conservative 172; Mismatches 290; Indels 119; Gaps 18;

Qy 1 MSPNNQNEYIILDASSSTVSNSVRYPLANDOTTTLQNNYKDYLRMSGEPNLPFGNP 60
Db 1 MSPNNQNEYIILDASSSTVSNSVRYPLANDOTTTLQNNYKDYLRMSGEPNLPFGNP 60
Qy 61 EFTFIS-SSTVOTGIGIVGVLGALGVPPAGQIASFSYFIVGQLWPSSTVSWMIMKQVE 119
Db 61 EFTFIS-SSTVOTGIGIVGVLGALGVPPAGQIASFSYFIVGQLWPSSTVSWMIMKQVE 119
Qy 61 EVLVSGQDAKAAIDIVGKLLGLGVFPVGPVIVSLYTLQIDILWPSGKSEQMBIFMEQVE 120
Db 61 EVLVSGQDAKAAIDIVGKLLGLGVFPVGPVIVSLYTLQIDILWPSGKSEQMBIFMEQVE 120
Qy 120 DLIDOKITDSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVYIALELDF 179
Db 120 DLIDOKITDSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVYIALELDF 179
Qy 121 ELINQKIAETARNKALSELGLGNVYQLYLTALBEEENPNGSRALRDVRRFEILDSLP 180
Db 121 ELINQKIAETARNKALSELGLGNVYQLYLTALBEEENPNGSRALRDVRRFEILDSLP 180
Qy 180 VAKIPSFASISQEVPLLSVYAAQANLHLLLRDASIFGAEWGFTPGIBISTFYDQVTRTA 239
Db 180 VAKIPSFASISQEVPLLSVYAAQANLHLLLRDASIFGAEWGFTPGIBISTFYDQVTRTA 239
Qy 181 TQWPSFRVNTFVFPFLTVYMAANLHLLLRDASIFGEENGWSTTTNNYIDRQMKLTA 240
Db 181 TQWPSFRVNTFVFPFLTVYMAANLHLLLRDASIFGEENGWSTTTNNYIDRQMKLTA 240
Qy 240 QYSDYCVKNTGTLGDKLKTNAASWLKYHOFREMLLLVLDVALPNYDTRTYPIETTA 299
Db 240 QYSDYCVKNTGTLGDKLKTNAASWLKYHOFREMLLLVLDVALPNYDTRTYPIETTA 299
Qy 241 EYSDHCYKWTETGLAKLGTSAKQWYDYNQFRREMLAVLDVVALPNYDTRTYPIETRA 300
Db 241 EYSDHCYKWTETGLAKLGTSAKQWYDYNQFRREMLAVLDVVALPNYDTRTYPIETRA 300
Qy 300 QLTREVYTDPIVFNRETSRGCRRWSNSDISPSEVESAVIRSPHLDILSEIEFTYTRA 359
Db 300 QLTREVYTDPIVFNRETSRGCRRWSNSDISPSEVESAVIRSPHLDILSEIEFTYTRA 359
Qy 301 QLTREVYTDPLGANVSSIG--SW-YDKAPSGFVIESSVIRPHFDYITGLTVYTSQR 356
Db 301 QLTREVYTDPLGANVSSIG--SW-YDKAPSGFVIESSVIRPHFDYITGLTVYTSQR 356
Qy 360 GLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIYKDYLANDKDIFQVRSIGA 418
Db 360 GLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIYKDYLANDKDIFQVRSIGA 418
Qy 357 S--ISSARIYRWAGHQISYHRVSRGSLQOMYGTQNLHSTSTFDTNVDIVKTLSDKA 414
Db 357 S--ISSARIYRWAGHQISYHRVSRGSLQOMYGTQNLHSTSTFDTNVDIVKTLSDKA 414
Qy 419 DLAN-----YAAVYGVYASFTLLDK--NTGSGSGVGYTSRPHHTMQVCTQNYTID- 470
Db 419 DLAN-----YAAVYGVYASFTLLDK--NTGSGSGVGYTSRPHHTMQVCTQNYTID- 470
Qy 415 VLADIVVPGYTYIFPGAPVEFEFFWVNLNTRK-----TLKYNPVSK 456
Db 415 VLADIVVPGYTYIFPGAPVEFEFFWVNLNTRK-----TLKYNPVSK 456
Qy 471 -----ELPPE--NEPLSRGYSRHLSHITSYSFSKNASSPARYGNLPFAWTHRS 517
Db 471 -----ELPPE--NEPLSRGYSRHLSHITSYSFSKNASSPARYGNLPFAWTHRS 517
Qy 457 DIIASTRDSLELPPETSDQPNYESYSHRICHITSIPATGNTI-----GLVPFVSWTHRS 511
Db 457 DIIASTRDSLELPPETSDQPNYESYSHRICHITSIPATGNTI-----GLVPFVSWTHRS 511
Qy 518 ADVNTVYSKTIQIPVVKAAHTLVSGTIVKGRGFTGCGNLIKRTSS-----GPLAYTSVSV 573
Db 518 ADVNTVYSKTIQIPVVKAAHTLVSGTIVKGRGFTGCGNLIKRTSS-----GPLAYTSVSV 573
Qy 512 ADLNTIYSKTIQIPVVKAWDNLFPVVPVVKGRGHTGCDLLQYNRSVGTGLFLARYGL 571
Db 512 ADLNTIYSKTIQIPVVKAWDNLFPVVPVVKGRGHTGCDLLQYNRSVGTGLFLARYGL 571
Qy 574 KSPLSQRYRARIYASTNRLRFLVTSIGTRIYSINVKNTWKNGDDLTFTNFDLATIGTAP 633
Db 574 KSPLSQRYRARIYASTNRLRFLVTSIGTRIYSINVKNTWKNGDDLTFTNFDLATIGTAP 633
Qy 572 ALEKAGKRYRVLRYATDADIVLVHND-----QIQMPKTNWPGEDLTSKTFKVADAITTL 626
Db 572 ALEKAGKRYRVLRYATDADIVLVHND-----QIQMPKTNWPGEDLTSKTFKVADAITTL 626
Qy 634 ---TFPNVSDSLTVGADSPAS-CGEVVYVDKFLIPVNATFEAREDDLOVAKAVNGLFTSK 689
Db 634 ---TFPNVSDSLTVGADSPAS-CGEVVYVDKFLIPVNATFEAREDDLOVAKAVNGLFTSK 689
Qy 627 NLATDTSALAKHNLGEPDNTSLGIVYVDRIEPIFVDTEYREQDLEAAKAVNALFTNT 686
Db 627 NLATDTSALAKHNLGEPDNTSLGIVYVDRIEPIFVDTEYREQDLEAAKAVNALFTNT 686
Qy 690 KDALQTSVTVQVNOAANLVECLSDLELYPNEKRLMDAVKEAKELVOARNLLQDTGNRI 749
Db 690 KDALQTSVTVQVNOAANLVECLSDLELYPNEKRLMDAVKEAKELVOARNLLQDTGNRI 749
Qy 687 KDGLRPGVTDYEVNOAANLVECLSDLELYPNEKRLLPDAVREAKELSEARNLLQDPDQEI 746
Db 687 KDGLRPGVTDYEVNOAANLVECLSDLELYPNEKRLLPDAVREAKELSEARNLLQDPDQEI 746
Qy 750 NGENGWGTSIGIEVABGCVLPKORSRLRLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 809
Db 750 NGENGWGTSIGIEVABGCVLPKORSRLRLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 809
Qy 747 NGENGWGTSIGIEVABGCVLPKORSRLRLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 806
Db 747 NGENGWGTSIGIEVABGCVLPKORSRLRLTSAREIDTETPTLYLQQIDESLLKPYTRYKL 806
Qy 810 KGFIGSSQDLIELKIRHANCQIVKRVNPNLLPDVLVUNSCGIDRCSEQQYVDANTALEN 869
Db 810 KGFIGSSQDLIELKIRHANCQIVKRVNPNLLPDVLVUNSCGIDRCSEQQYVDANTALEN 869
Qy 807 RGFVSSQGLEIETIRHQTRNRIKVNVPDDLPLPDVSPVNSDGSINRCSQKYVNSRLEVEN 866
Db 807 RGFVSSQGLEIETIRHQTRNRIKVNVPDDLPLPDVSPVNSDGSINRCSQKYVNSRLEVEN 866
```


RESULT 10

US-10-614-524-6
; Sequence 6, Application US/10614524
; Publication No. US20040016020A1
; GENERAL INFORMATION:
; APPLICANT: Arnaut, Greta
; APPLICANT: Boets, Anemie
; APPLICANT: Damme, Nicole
; APPLICANT: Mathieu, Eva
; APPLICANT: Vanneste, Stijn
; APPLICANT: Van rie, Jeroen
; TITLE OF INVENTION: Insecticidal proteins from *Bacillus thuringiensis*.
; FILE REFERENCE: NEWBTSUS2
; CURRENT APPLICATION NUMBER: US/10/614,524
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US/09/739,243
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/173387
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: *Bacillus thuringiensis*
US-10-614-524-6

Query Match 47.78; Score 2884.5; DB 15; Length 1152;
Best Local Similarity 49.14; Pred. No. 3.8e-267;
Matches 578; Conservative 200; Mismatches 365; Indels 33; Gaps 14;

QY 1 MSPNNQVEYILDASSSTVSNSVRYPLANDQTTTLQNNNYKDYLRMSGCEPNELFGNP 60
DB 1 MNRNNDYEVIDASNCGCASDDVQVPLARDPNAVFQNMHYKDYLTQYDGYDTSFNP 60

QY 61 ETPISSSTV-QTGIGIVGVGLGALGVFPAGIASFYSFIVGQLWPSSTVSWMIMQVE 119
DB 61 NLSINPRDVLQGTGINIVGRLLGFLGVFPAGOLVFTYFTLLNQLWPTTNDNAVWEAFMAQIE 120

QY 120 DLIDOKITDSVRKTAGLGLDGLDVYOKLSKNMLNENDTFRASVVVTVYIALRDP 179
DB 121 ELINQRISEAVGTAADHLTGLHDNYELVEALSEWLERPNAAT-NLLFNRTFLDLSLF 179

QY 180 VAKIPSPA1-----SGQVPLLSVYAQAANLHLLLRDASIFGAEGFTPGESTFYDRQV 235
DB 180 TQMPSPFTGPGSQNYAVPLTVYAQAANLHLLLRDASIFGAEGFTPGESTFYDRQV 239

QY 236 TRTAQYSDYCVKYNVTGLDKLGTNAASWLKYHOFREMTLLLDLVALPNNYDTRTPI 295
DB 240 ERTQYTNHCVTYNTGLDRGTNTESWLNHYHFRREMTLWMDLVALPNNYDTRTPI 299

QY 296 ETQAOLTRVYTDIVFNRETSQFCRWSLNSDISSEVESAVIRSPHLDLISEIFY 355
DB 300 GANPOLTREITDPIVYNPANQICRRWGNPNYTFSELENAPFIRPHLFDRLNRTIS 359

QY 356 TTRAGLPLNNTLEYLVWVGHISIKYKNTNASSALERNYGTITSNKIKYVDLANKDIFQVRS 415
DB 360 RNRYPATPNS-YLDYWSGHLLQYANNPTTYSYQITSN-TRLFNTN-CANAIDS 416

QY 416 LGADLANVYQVGVPSFTLLDKNTGSGVGGFTYSKPHTTMOCVTQNYNTIDEIPPE 475
DB 417 RARNFGLNLYANLYGVSYLNI-----PPTGVMSBIT-SAPNT---CWQDLTTTEELPLV 465

QY 476 NEPLSRGSHRLSHITSYSPSKNASSP-ARYGNLPVFAWTHRSADVTNVTYSKTIQIPV 534
DB 466 NNPN-----NLLSVTLFRLNTTQGGPLATVGFVPTVTRQDQVDFNNIITPNTIPIV 520

QY 535 VKARTLVSGTIVKPGFTGNNILKRTSSGGLAVTSVSKPLSQRYRARIYASTNLR 594
DB 521 VKAYELSSGATVWVGPGTGGDVIRRTNTGFGAIRSVTGLTQRTIRPRIYASTIDPD 580

QY 595 LFTVTSSTRIYSINVAKTNKGGDLTFTNTFDLATIGTAFSTFNSYSDSLTVGADSFASGGE 654
DB 581 FFVTRGTTINNFRFTNMGRQESRYESRTVEFTTPTFQSDIIRTSIQLSGNGE 640

QY 655 VYVDKFEELIPVNATFRAEEDLDVAKAVNGLFTSKDALQTSVTDYQVQANLVCELSLD 714
DB 641 VYLDREIIPVNPTRAEEDLEAAKAVASIFTRDGLQVNVTDYQVQANLVCELSLD 700

QY 715 ELYPNERGLMWDVAKBAKRLVQARNLLQDQFNRING--ENGWGTSGTGIEVAEGVLPKD 772
DB 701 EQYAHDKMLLEAVRAAKRLSRERLLQDPDFTNTINSTENGWTKASNGVTISEGGPFYKG 760

QY 773 RSLRLTSAREIDTETPTLYLQQIDESLILKDYTRYKLKGPFGSSQDLLEIKLIRHRANQIV 832
DB 761 RALQLASAR----ENYPTIYQKVDASELKPTRYRLDGFVKSSQDLLEIDLIHHRKHLV 816

QY 833 KNPVDMLLPDVLPVNSCGIDRCSEQQYVDANLALENNGB--NGNNSSDSHAFSPHIDTGE 891
DB 817 KNVLDNLVSDTYDDSCSGINRCCEQQMVAQLETEHHHPMDCCCEAAQTHFSSVIDTGD 876

QY 892 IDLANENTGIWVPKIPPTINGVATLGNLELVEBGLSGETTLERAAQQOQOQDKMARKEGA 951
DB 877 LNSTVDQGIWIPKAVRTTDDGYATLGNLELVEBGLSGETTLERAAQQOQOQDKMARKEGA 936

QY 952 SEKAYYAAKQAIIDRLPADYODOKLNSGVHSDMLAAQNLVQSIPYVYNDALPRIGMNYT 1011
DB 937 TDRVYQDAKQSIHNLFDYDQQLNPQIGHADIMDAQNLVASISDVYSDAVLQIPGINYE 996

QY 1012 SPTELTNRLQOANLVYDLRNAIPNGDFRNLGNLSDNMTSDVNVQQLSDTSVLVIPNNSQV 1071
DB 997 IYIELSNRLQOASVLYTSRNVQNGDFNGLSDSNATAGASVQDGNTHFLVLSHWDQV 1056

QY 1072 SQQFTVQPNRYVYLVRTARKEGVGDGYVIRIDGANQOTETLFTNTCDDDTGVLSDAQTYSI 1131
DB 1057 SQQFVQPNCKYVLRVTAEBKVGSGDGYVTRIDGAHHTETLTFNACDYDINGTYVTDNTYL 1116

QY 1132 TKTVETPTSEQVWIDMSSETGVNIESVELVLEE 1167
DB 1117 TKCVIFYSHTEHMMVYEVNTEGAFHIDSIEPVETER 1152

RESULT 11
US-10-099-285-74
; Sequence 74, Application US/10099285
; Publication No. US20030105319A1
; GENERAL INFORMATION:
; APPLICANT: Schnepf, H. Ernest
; Narva, Kenneth E.
; Walz, Michelle
; Stockhoff, Brian
; Muller-Cohn, Judy
; TITLE OF INVENTION: Toxins Active Against Pests
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/099,285
; FILING DATE: 15-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/002,285
; FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 1150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-099-285-74

Query Match 46.7%; Score 2825.5; DB 14; Length 1150;
Best Local Similarity 48.6%; Pred. No. 2,6e-222;
Matches 574; Conservative 197; Mismatches 364; Indels 47; Gaps 17;
QY 1 MSPNNQNEYILDASSSTVSNSVRYELANDQTTLLQNMNYKDYLRMSGE-----NP 54
DB 1 MNRNNEVEIIDAFCPCPSDDDDVRYPLASDPAAFQNMNRYLYQYDGDYTGSLNP 60
QY 55 ELFGPPTFSSSTVGTGIGIVGVLGALGPPFAGQIASFSYFVQGLWPSSTVSWEI 114
DB 61 NLSINPR-----DVLQGTINIVGRILGFLGVPFAGQLVTFYTFLLNQLMPTNDNAWEAF 115
QY 115 MKQVEDLIDOKITDSVRKTAGLQGLGDVQKSLKWLLENRNDTRARSVVVTVIA 174
DB 116 MAQIEELIDKISAQVNRNLDLGLHDYEEYLALEWLERPENGARA-NLVTQREN 174
QY 175 LEQDPVAKIPSPAIS-GQE---VPLLSVYAQAANLHLLLRDASIPGAEWGTFGEISTF 230
DB 175 LHAPVTRMPSFGTGPQSDQDAVALLVYAQAANLHLLLRDARIYGARWGLQCGQINLY 234
QY 231 YDRQVTRTAQYSDYCKVNNGLDKLGTNAASMLKYHOFREMTLLVLDLVALPPNYDT 290
DB 235 FNAQQRTRITYNHCVETINRGLDVRGTNTEMLNHYRFRREMTLMDLVALPFPYNV 294
QY 291 RTYPIETTAQTRKYVTDPIVFNRETSGFCRRWSLNSDISFSVESAVIRSPHLDLIS 350
DB 295 ROYPNGANPQLTREIYTDPIVNPANQICERWGNPNYTFSELENAFIRPHLPERLN 354
QY 351 EIEFYTTRAGLPLNNTEYLEYVWGHISIKYKTNASSALERNYGTITSNKIKYDLANKDI 410
DB 355 RLITSRNRYTAPTNS-FLDYWSGHTLQSQHANNPTVETSYGQITSN-TRLENTTN-GA 411
QY 411 FQVSLGADLANVYAQVGVYASFTILDVNTGSSGVGFTYSKPHITMQVCTQNYTID 470
DB 412 RAIDSRAENFGNLYANLYGV--SSLNIPP--TG-----VMSEITNAANTCRODLTTE 460
QY 471 EIPPNEPLSRGSHRLSHITSYSFSPKSNASP-ARYGNLPPFAWTHRSADVNTVYSDKI 529
DB 461 ELPLENNP-----NLLSHVFLRNTTQGGPLATLGFVPPVVTREDVDFNITADRI 515
QY 530 TQIPVVKATLVGTTVIGKFGTGNILKRTSSGPLAYTSYVKSPLSQRYRARIYAS 589
DB 516 TQLPVWKASEIGGTTVVKGFGTGDILRRTDGGAVGTIRANVAPLTPQYRILRYAS 575
QY 590 TTN--LRLPVTISGRTRYISINVKNTMKGDDLTFTFPDLAIGTAPTNSDSLTVCAD 647
DB 576 TTSFVNLFVNNSAA---GFTLPSTMAQNGSLTYESFNTLEVTHTIRFSQSDTTLRLNIF 632
QY 648 SPASGEVYVDKFLIPVNATPEAEEDLDVAKAVNGLPTSKDALQTSVTDYOVNOAAN 707
DB 633 PSISQGEVYVDKLEIVIPNPTREAEEDLEDAKAVASLFTTRDGLQVNTDYQVDOAAN 692

QY 708 LVBCLSDBLYPNKGMKMLWDAVKEAKRLVQARNLLQDTGFNRING--ENGWTCSTGIEVAE 765
DB 693 LVSCLSDBQYGHDKMMLLEAVRAAKRLSRERLLQDDPDFNEINSTEENGWAKASNGVTISE 752
QY 766 GDVLPKDRSLRLTSAREIDTETPTVLYQOIDESESLKPYTRKLGKFGIGSSODLEIKLIR 825
DB 753 GGPFFKGRALQASAR-----ENYPYIYQKVDASTLKPYTRKLGKFGVQSSQDLEIDLH 808
QY 826 FRANQIVKXVPDNLPLPVLVNSCGGIDRCSEQQVYDANLALENNENGNMSSDSHAFSP 885
DB 809 EHKVHLVKNVPDNLVSDTYSDGSCSGINRCBQHQVDVQLDAEDHPKDCCEAAQTHEPSS 868
QY 886 HIDTGEIDLNTGTVVVKIPPTNGYATLGNLELVBESGPLSCETLRAQQOQOQOQDKM 945
DB 869 YIHTGDLNASVDQGIWVVLQVTRTDDGYATLGNLELVBESGPLSCETLRAQQOQOQOQDKM 928
QY 946 ARKEGASEKAYAAKQAIIDRLFADYQDOKLNSGVMSDMLAAQNLVQSPYVYNDALPEI 1005
DB 929 GRKAETDRIYQDAKQAINHLFVYQDQOLQSPFVGMAIIDAQNLIASISDVYSDAVLQI 988
QY 1006 FGMNYSFTSLTNRLQQAANLYDLRNALPNSGDFRNLGLSDMNATSDVNVQQLSDTSVLVIP 1065
DB 989 FGINVEMTSLNRLQQAASLYTSRNVVQNGDFNSGLDSMNATTDVAVQDQGNMHPVLIS 1048
QY 1066 MNSQVSOQFVQPNRYVLRVARTAKSGVGGYVLRDQANQETETLTNLCDDDTGVLSA 1125
DB 1049 HMDAQVSOQFVQPNCKYVLRVARTAKSGVGGYVLRDQANQETETLTNLCDDDTGVLSA 1108
QY 1126 DQTSYITKVTFTSTROVWIDMSSETGVFNIESVELVLEBS 1167
DB 1109 NDNSTIKELVFPYKTEHMMWVEVSETEGTFTYDSDIEFTQOE 1150
RESULT 12
US-10-099-285-76
; Sequence 76, Application US/10099285
; Publication No. US20030105319A1
; GENERAL INFORMATION:
; APPLICANT: Schnepf, H. Ernest
; Wicker, Carol
; Narva, Kenneth E.
; Walz, Michelle
; Stockhoff, Brian
; Muller-Cohn, Judy
; TITLE OF INVENTION: Toxins Active Against Pests
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/099,285
; FILING DATE: 15-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/002,285
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/886,615
; FILING DATE: 1-JUL-1997
; APPLICATION NUMBER: US 08/674,002
; FILING DATE: 1-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-701C2

TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 1134 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 76:
US-10-099-285-76

Query Match 44.8%; Score 2710; DB 14; Length 1134;
Best Local Similarity 48.1%; Pred. No. 7.9e-213;
Matches 559; Conservative 195; Mismatches 345; Indels 62; Gaps 19;

Qy 39 NMNYKDYLRMSEGENPEL-----FGNPETFISSSTVGTGIGVQVILGALGVPAGQIA 92
Db 4 NPNINECIPYCLSNPEVEVLGGERGNVT-----GLQTGIDIVAVVVVGLGPPVGGILT 58

Qy 93 SFYSPVQLWPSSTVSVMIMKQVBLIDQKITDSVRKKTALAGLGLDGLDVIYKSL 152
Db 59 GFLSTLFGFLWPSNDQAWAEPIEQMBELIEQRISDQVTRTALDDLTGIGNYNYQYLI 118

Qy 153 KNWLENRNDTRARSVVVYQYIALELDFVAKIPSPAISGO-----EVPILLSVYQAANLHL 207
Db 119 KEWEERPNGVRA-NLVQLRFBILHALFVSSMPSPFG-SGFGSGRFOAQLLVVYQAANLHL 176

Qy 208 LLLDASIFGAEWGFTPEISTFTYDROV-TTAQYSDYCVKWTGTLKLGKGTWASWLK 266
Db 177 LLLADAKEYGARWLGRSIOGNLYFNELOTRDYTNHCVNNGLAGLGRGTSABSWLK 236

Qy 267 YHQFRREMTLLDLVALFPNYDRTYPIETTAQLTREVVYDPIVFNRETSGP----CR 322
Db 237 YHQFRREMTLLDLVALFPNYDRTYPIETTAQLTREVVYDPIVFNRETSGP----CR 322

Qy 323 RWSLNSDISFEVSASVIRSHPLFDLSEIFBYTTRAGLPLNTEYLEYVWVGHST--KYK 380
Db 297 RWOQTSAMTFGNLENAILSSPHLFDLSEIFBYTTRAGLPLNTEYLEYVWVGHST--KYK 380

Qy 381 NTNASSALERNYGITNSKIKYDLDANKDILQVRS---LGADLANVYQVYGVVPVASF 437
Db 356 ASGPTTLVRRNYGSGTTS-IVNYFSPNDRDVOQINRSHTGILGFQV--APLFGITRAQF-- 410

Qy 438 LDKNTGSGVGGFTYKPHTTNQVCTQNYNTIDEIP--PENEPISRGYSHRLSHITSYSP 495
Db 411 -----YFGIYSVTQRNALTCEQYNSIDELPSLDNPEPISRSYSHRLSHITSYLH 461

Qy 496 -----SKNASSPARYGNLVPFAMTHRSADVNTVYSDKITQIPVVKATHLVSGTIVYK 549
Db 462 RVLTIDGINIYS-----GNLPTVWVTHRDVLDLTNTITADRTQLPLVKSFBIPAGTIVVRG 517

Qy 550 PGFTGNNILKTSGLPATTYSVKSPISORVYRIRVASTNLRFLTISGTRIYSINV 609
Db 518 PGFTGGDILRRTGVGTFTTIRVTAPUTQYRIRFPASTNLPFGIRVGRDQVNYPDF 577

Qy 610 NKTWKNKGDLLTFNFDLATIGTAPTFNSYSLTVGADSPASGGEVYVDKPELIPVNA 669
Db 578 GRVWNRGDELAYESFATREFTTDFNFRQPELISVPANAFSAGQEVYFDRIEILIPVNP 637

Qy 670 EABEDLVAKVAVNGLFTSKKDALQTSVTDYQVNOQANLVCLSDLEYPNKRMLMDAVK 729
Db 638 EAKEDLEAAKAVASLFTTRDGLQVNVKDYQVQQAANLVCLSDLEQYGDYKQMLLEAVR 697

Qy 730 EAKLVQARNLLQDTGNRING--ENGWTCGTGLEVAEGDVLFDKRSRLTSAREIDTET 787
Db 698 AAKLSRERNLLQDPDFNTINSTEENGWAKSGVNTISEGPFYKGRALQLASAR-----EN 753

Qy 788 YPTLYQOQIDSLKPYTRYKLGFGIGSSQDLKLRHRANQIVKQVNPDLNLLPDVLPVN 847
Db 754 YPTTYQKVDASLKPTRYRSDGPFVKSQDLKLRHRANQIVKQVNPDLNLLPDVLPVN 813

Qy 848 SCGGIDRCSEQQYVDANLALENGEB-NGMSSDSHAPSHPHIDTCEIDLNTGNTIWWVPKI 906
Db 814 SCGINRCQEQQVNAQLETEHHHPMDCCBAAGTTFHFSSYIDTGLNLSVVDQGIWPKV 873

Qy 907 PTTNGVATLGNLLEVEBGLSGETILERAQQOQOQDKMKRKGASEKAYAAKQAI 966
Db 874 RTTDTATLGNLLEVEBGLSGETILERAQQOQOQDKMKRKGASEKAYAAKQAI 933

Qy 967 FADYQOQKLSGVMSDMLAQLVQSIPIVYVNDALPEIPGMNYTSFTLTLNRLQQA 1026
Db 934 FVDYQOQKLSGVMSDMLAQLVQSIPIVYVNDALPEIPGMNYTSFTLTLNRLQQA 993

Qy 1027 YDLRNLAIPIGDFRNLGSLDNATSDVNVQOLSTSVLVIPIWNSQVSOQFTVQNTY 1086
Db 994 YTSRNVQNGDFNGLSDWNAATAGASVQDQGNTHFLVLSHMDAQVQQFRVPQNC 1053

Qy 1087 VTARKEGVGDGVVIRDCANOTETITFNICDDDTGVLSDAQTSYTKTVEFTPTST 1146
Db 1054 VTAKEVGGDGVVIRDCANOTETITFNICDDDTGVLSDAQTSYTKTVEFTPTST 1113

Qy 1147 DMSETEGVFNIESVELVLEEE 1167
Db 1114 EWNTEGAFHIDSIEFVETEK 1134

RESULT 13
US-09-988-462-7
; Sequence 7, Application US/09988462
; Publication No. US20030046726A1
; GENERAL INFORMATION:
; APPLICANT: Kozel, Michael G.
; Desai, Malini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Byola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Syngenta Biotechnology, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09988,462
; FILING DATE: 20-NO. US20030046726A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/547,422
; FILING DATE: 11-APR-2000
; APPLICATION NUMBER: US 08/459,504
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-188051
; TELECOMMUNICATION INFORMATION:

```

; ; TELEPHONE: (919)541-8587
; ; TELFAX: (919)541-8689
; ; INFORMATION FOR SEQ ID NO: 7:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 1207 amino acids
; ; TYPE: amino acid
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: protein
; ; SEQUENCE DESCRIPTION: SEQ ID
US -09-988-462-7

```

Query Match	41.7%	Score 2519.5	DB 10	Length 1207
Best local Similarity	44.3%	Pred. No. 3.9e-197		
Matches 548	Conservative 178	Mismatches 372	Indels 139	Gaps 20
QY	37	LQNNYKDYLRMSGCEPNELFGNPETFISSSTVGTGIGVQVIGALGVPPAGQIAFPYS	96	
DB	4	LPDARIEDSLCIAGS-----NNIDPFVSASTVGTGINIAGRIILGVLPFGPAGOLASFYS	57	
QY	97	FIVGQLPSSSTVSVMEMIKQVEDLIQDKITDSVRKTALAGLQGLGDLVYQKSLKNWL	156	
DB	58	FLVGELWPRGR-DQWEIFLEHVEQLINQIITENARNTALARIQLGLGDSFRAYQQSLDNL	116	
QY	157	ENRNDTRARSVVTOYTALELDVFAKIPSPFAISQGEVPLLSVYAQAANHLILLARDASIP	216	
DB	117	ENRDDARTSVLYTOYTALELDVFLNAPLFAIRNQEVPLLMVYAQAANHLILLARDASLF	176	
QY	217	GAEMGFTPGELSTVDQVTRTAQVSDYCVKMYNTGLDKLGTNAASLKYHQPRREMTL	276	
DB	177	GSEPLTSQELQRYEYQVETRYSDYCVENYNTGLSLAGTNAASWVRNQFRDLTL	236	
QY	277	LVLDLVALPNYDTRTYPIETTAQLTRVYTDPIVFNRETSGGFCRMSLNSDITSPSEVE	336	
DB	237	GVLDLVALPNSYDTRTYPIINTSAQLTRVYTDALGATGVNMASM--NMYNNAPSPSAIE	294	
QY	337	SAVIRSHPLFDILSEIFPYTTRAGLPLNNTLEYLVWCHSIKYK-----NTNASSALER	390	
DB	295	AAAIRSHLLDFLEQLTFS--ASRWSNTKMTYWRGHTIQSRIPOGGLMTSTHGATNT	352	
QY	391	NYGTITSNKIKYYDLANKDIPQVRSAGADL--ANYAAQVYGVVPYASFTLLDRNKGSGSVG	448	
DB	353	SINPVTLR-----PASRDVYRTESYAGVLLGWILEPIHGVPVTRPFTNPQWISDR-G	405	
QY	449	GFTYSKPHHTMQVCTQNTYNTIDEIPPE--NEPLSRGYSHRLSHITSYSPSKNASSPARYG	506	
DB	406	TANYSQPYSPGLQKDSCT--ELPPEPTTERPNYESYSHRLSHITGIIQSR-----V	455	
QY	507	NLPVFAWTHRSADVNTVYSBKITQIPVVKAHTLVSGTTVYIKGPGFTGNNILKRTSSGPL	566	
DB	456	NVPVYSWTHRSADRTNIGPNRIQIPWKASBELPQGTTVVRGPGFTGDIILRNTTGGF	515	
QY	567	AYTSVSVKSPISQRYARIRYASTNLLPVTISGTRIYSINVNKTWNKGGDILTFNTFDL	626	
DB	516	GPIRVTVNGPLTQRYRIGRFYASTVDFPVSRGCTTVNNFRFTTNNSGDELXYGNFVR	575	
QY	627	ATIGTATFYSNYSLSLTGADS PASGCEVYVDKELIPVNATPFAEEDLDVAKAVNCLF	686	
DB	576	RAFTPTPTFTQIDIIFTSIQGLGNGEVYIDKLEIIPVTAFEEABYDLERAOASVNALF	635	
QY	687	TSKXD-ALQTSVTDVQVQAANVECLSDLELYNEKRLMDAVKEAKELVQARNILLQDTG	745	
DB	636	TNTNPRRLKTDVTDYHIDOVSNVACLSDLEFCLDEKRELLEKVKYAKELSDERNLLQDPN	695	
QY	746	FNRIN-----GEMWTGSGTGIEVARGDVLFKORSRLBLTSAREIDT	785	
DB	696	FTSINKOPDFTISTNEQSNFTSIHQSEHGWSGSENIITIQEGNDVPFKENYVTLPGT---	752	
QY	786	ETYPYLYQQIDESILKPYRYKLGKPGTSSODLEIKLIRRAQIVKQNVF--DNLLPDV	843	
DB	753	ECYPYLYQKIGESLKYATRYQLNGYIEDSQDLEIYLIRYNAKHETLDVPGTSLSLWPLS	812	
QY	844	L--PVNSCGGIDRCSEQQYVDANTALENNGBNG--NMSSDSHAFSPHIDTGTBIDLNTGTI	900	

```

Db      813  VESPIRGCGPNRCA--PHEWNPDLDCSRDGEKCAHHSHHSFLDIDVCCDTLHMLGV 870
Qy      901  WVVPKIPITTINGYATLGNLELVBSPISGETLERAQOOQOQOMARKGASEKAYYAAK 960
Db      871  WVVPKIKTQEGHARLGNLEPIBEKPLLGALSRVKAEGKWRDKREKLOLETKEVYTEAK 930
Qy      961  QATDELPAFYQDQKLNSGVEMSDMLAAQNLVQSPIPYVYNDALPEIPGMNTYSTPELTNRL 1020
Db      931  EAVDALPVDYSQYORLQADTWIGMIHAADKLVHRIRAYLSELPIPOVNAEIPFEELECHI 990
Qy      1021  QOAWNLVDLENAIPNGDFRNLGSLDMNATSDVNVQQLSDTSVLVIPNMNSOVSOQPTVPQN 1080
Db      991  ITAISLDYARNVKNRGDFPNGLTCMNVKGVHDVQQSHRSSDLVIPWEAREVSQAVRVCPG 1050
Qy      1081  YRVVLRTAKRGVGDGVYIIRGANTQETLT- 1113
Db      1051  CGYILRTVAYKEGVEGCVTIHEIENNTDELKPNREBEVEYPTDTGTCNDYTAHQGTAG 1110
Qy      1114  -----NICDDDTGVLSAD--QTSY 1130
Db      1111  CADACNSRNAGYDAYEVDVTASVNVKPTVEETTYTDRDNHCEYDRGVNTPVPVPGY 1170
Qy      1131  ITKTVEFTPTBOVWIDMSSETGVNIESVELVLEE 1167
Db      1171  VTKELEFPPTDTWIEIGTEGKPIVDSEVLLAMEE 1207

RESULT 14
US-10-428-961-63
; Sequence 63, Application US/10428961
; Publication No. US2003023711A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rugar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; TITLE OF INVENTION: Polynucleotides, Compositions, and Methods of Use (Amen
; FILE REFERENCE: MESCO201--1
; CURRENT APPLICATION NUMBER: US/10/428,961
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 63
; LENGTH: 1227
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-428-961-63

```

Query Match	40.7%	Score 2457;	DB 15;	Length 1227;
Best Local Similarity	42.8%;	Pred. No. 5.4e-192;		
Matches 541;	Conservative 197;	Mismatches 392;	Indels 134;	Gaps 20;
QY	1	MSPNNQKEYETLDDASSTSTSDNSVRYPLANDQTTTLQNNNYKYDLRMSGCEPNELFGNP	60	
DB	1	LTSNRKNEEINLALSTPAVNSHSAQNWLTSD-----ARIEDSLCIAEG-----NNI	47	
QY	61	ETPISSTVTQGIIGVQVILGALGVPPAGQIASFYPIVQGLMPSSTVSVMHMKQVED	120	
DB	48	DPFVSASTVTQGINIAGRIILGVLGVPAGQIASFYSELVGELMPGRG-DPWEIPLFLEHVEH	106	
QY	121	LIDQKITDSVRKATAGLQGLGDLGVYQKSLKNMLNRMNDTARSVVVVYQYIALELDFV	180	
DB	107	LIRQVTEINTDATALRQLQGLNSPRAYQQSLDMLNRRDDATRSVLYTYQYIALELDFL	166	
QY	181	AKIPSPAI SQGVPELLSVYAAQANHLHLLLRDASI PCAENGFPTGEISTSTPYDQVTRTAQ	240	
DB	167	NAMFLPAIRQGVPELLMYYAAQANHLHLLLRDASLFGSEFLGTSQETQRYERQVSKTRTE	226	

```
Qy 241 YSDYCVKWTGTLGDKLGTNAASWLVKQHFQREMTLLVLDLVALFPNRYDRTYPIETTAQ 300
Db 227 YSDYCARWYNTGLANLGTNAESWLVYQNFREDLTGLVLDLVALFPNRYDRTYPIETTAQ 286
Qy 301 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHPLFDILSBIEFYTTTAAQ 360
Db 287 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHPLFDILSBIEFYTTTAAQ 345
Qy 361 LPLANTVEYVWGHSHIKYKNTNASSALERNYGTITSNKIKYDILANKDIPQVSLGADL 420
Db 346 -RWSNTQYMWYVGHSHIRSLTSGTSTHGTNTSINPVLTQFTSRDVTYTESFAGIN 404
Qy 421 ANYYAQVGVYVYASFTLLDKNKTGSGVGGFTYKSPHTMQCTQNYNTIDIPPE--NEP 478
Db 405 ILLTPVNGVFWARFNM--RNPLNSLGRSLLYTIGYT--GVGTQLFDSLETPEPTTERP 460
Qy 479 LSRGYSHRLSHITSYSPSKNASSPARVGN---LVPFANTHRSADVNTVYSDKIQIPLV 535
Db 461 NYESYSHRLSHIRLIS-----GNTLRAPVYSWTHRSADRTNTISSDITQIPLV 509
Qy 536 KAHTLVSGTIVIKGPGFTGGMILKRTSGGPLAYTSVSVKSPLOKRYARIRYASTNLR 595
Db 510 KSFMLNGSTSVSGPGFTGGDIIRTNVNGSVLSGLNFPNNTSLQRYRVRVRIASQTMVL 569
Qy 596 FVTISGTRIYSINVKNTMKGDDLTFTNTFDLTATIGTFTFNSYSDSLTVGADSPASGGEV 655
Db 570 RVTYVGGSTPDQGPSTMSANESLTSQSPFAEPVPGISASG--SOTAGISISNAGRTF 628
Qy 714 YVDFKELIPVNAATEAREDLVAKVANGLTSSKQ--ALQTSVTDYQVQNAANLVECLSD 714
Db 629 HFDKIEPIPTATPEABYDLERAQEAVALPTNTNPRLLKTGVTDYHIDEVSLVACLSD 688
Qy 754 ELYPNEKRLMDVAKEAKLVQANLLQDTGFRIN-----GNG 754
Db 689 EPLDREKRELBKVKYAKRLSDERLLQDPNFTSINKQDPNNSNROSNTSIIHQSEHG 748
Qy 814 WTGSGTEVAGBDVLFKDRSLRLTSAREIDTETPTLYQOIDSLLKPYTRYKLGPIG 814
Db 749 WNGSENITIQEGNDVFXYVTLPGT---FNECPTVLYQKIGABLKAVTRYQLSGVIB 805
Qy 870 SSQBLEIKLIRHRAQIVKQNP--DNLLPVL--PVNSCGGIDRCSEQQVVDANALENN 870
Db 806 DSQBLEIYLIRYNAKHETLDVPGTESVWPLSVESPIGRGSPNRCA--PHEWNPDLDCS 863
Qy 929 GENG-NWSSDSHAPSFIHDTGEIDLNENTGTWVYFKIPTTNGYATIGNLLEVERGSLGB 929
Db 864 CRDGEKCAHSHHSLDIDVCGTIDHENLGVVWVFKITQEGHARLGNLEFIEKPLIGE 923
Qy 989 TLERAQOQOQOQOQOQMARKGASKEYAAKQAIIDLRFADYQDQKLSGVMSDMLAON 989
Db 924 ALSRVKRAEKWKDKREKLQLETKRVYTEAKEAVDALFVDSQYDRLQADTNIGMHAADK 983
Qy 1049 LVQSIPTVYNDALPEIPGMNTYSTELTNLQQAANLYDLNAILPNGDFRGLSDWATS 1049
Db 984 LVHRIEAYLSLSVIFGVNABIFEELEGRITITISLYDAENVVYKNGDFNGLACWVK 1043
Qy 1109 DVNVQQLSDTSVLVIPNNSOVSOQFTQPNRYVVLVYARKEGVGDCYVIRDQANQTE 1109
Db 1044 HVDVQSHHRSVLVPEWEAEVSQVRVCPGGRGILRVATYKGVGGCVTIHIEENNTD 1103
Qy 1129 TLTFNICDD-----DTGLVS-----ADQTS----- 1129
Db 1104 ELKPKNCEBSEVPTDGTCTNDYTAHQCTAVCNSRNAGYEDAVEVDVTTASVNYKPTYEE 1163
Qy 1130 -----YTKVTEFTPEEQWIDMSHTEGVNFIESVELV 1163
Db 1164 TYTDVRDNHCEYDYGYNVPPVAGYNTKLEYFPETDKVWIBIGETEGKFIVDSEVELL 1223
Qy 1164 LEE 1167
Db 1224 LME 1227
```

```
RESULT 15
US-09-826-660-23
; Sequence 23, Application US/09826660
; Patent No. US2001002690A1
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelmaun, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1186
; TYPE: EXT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Toxin encoded by synthetic B.t. gene
US-09-826-660-23
```

Query Match 40.6%; Score 2454.5; DB 9; Length 1186;

Best Local Similarity 43.1%; Pred. No. 8.3e-192;

Matches 536; Conservative 191; Mismatches 382; Indels 135; Gaps 19;

```
Qy 1 MSPNNEYETLDASSSTVSVDNSVRYPLANDQTTTLQNMNMYKDYLRMSGEENPELPGNP 60
Db 1 MTSRKNENEIINALSIPAVSNHSAQMWLSTD-----ARIEDSLCTAEG-----NNI 47
Qy 61 BTFISSSTVQTGIGIVGQVLAGALGVPFAGQIASPVSYFVIGQLMPSTSVSWEMIMKQVED 120
Db 48 DPFVSASTVQTGINIAGRILGVLGVPFAGQIASPVSYFVIGLMPGR--DPMEIFLEHVQ 106
Qy 121 LIDQKITSVRKLTALAGLOGLDGVYQKSLKNLENNDTRARSVVVTVQVIALELDFV 180
Db 107 LIROQVNTERTDALARLQGLGNSFRAYQOOSLKNLENDDARTSRVLYTVQVIALELDF 166
Qy 181 AKISFSAISQGVPLLSVYQAANLHLALLDASIFGAEWGTGCEISTFYDQVTRTAQ 240
Db 167 NAMPLPAIRNGEVPLLVYQAANLHLALLDASLFGSEFGLTSQEIORYYERQVKTRE 226
Qy 241 YSDYCVKWTGTLGDKLGTNAASWLVKQHFQREMTLLVLDLVALFPNRYDRTYPIETTAQ 300
Db 227 YSDYCARWYNTGLANLGTNAESWLVYQNFREDLTGLVLDLVALFPNRYDRTYPIETTAQ 286
Qy 301 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHPLFDILSBIEFYTTTAAQ 360
Db 287 LTRVYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSHPLFDILSBIEFYTTTAAQ 345
Qy 361 LPLANTVEYVWGHSHIKYKNTNASSALERNYGTITSNKIKYDILANKDIPQVSLGADL 420
Db 346 -RWSNTQYMWYVGHSHIRSLTSGTSTHGTNTSINPVLTQFTSRDVTYTESFAGIN 404
Qy 421 ANYYAQVGVYVYASFTLLDKNKTGSGVGGFTYKSPHTMQCTQNYNTIDIPPE--NEP 478
Db 405 ILLTPVNGVFWARFNM--RNPLNSLGRSLLYTIGYT--GVGTQLFDSLETPEPTTERP 460
Qy 479 LSRGYSHRLSHITSYSPSKNASSPARVGN---LVPFANTHRSADVNTVYSDKIQIPLV 535
Db 461 NYESYSHRLSHIRLIS-----GNTLRAPVYSWTHRSADRTNTISSDITQIPLV 509
Qy 536 KAHTLVSGTIVIKGPGFTGGMILKRTSGGPLAYTSVSVKSPLOKRYARIRYASTNLR 595
Db 510 KSFMLNGSTSVSGPGFTGGDIIRTNVNGSVLSGLNFPNNTSLQRYRVRVRIASQTMVL 569
```

```

Qy 596 FVTISGTRIYSINVNKTKNGDDLTFTNTPDLATICTAFTPNYSDSLTVGADSPASGGEV 655
Db 570 RVTVGGSFTFDQGFPSITMSANESLTSQSFRAEFVIGISASG-SQTAGISISNNAGRQTF 628
Qy 656 YVDKFEIIPVNATPEAREDLVAKKAVNGLFTSKD-ALQTSVTDYQVNOQANLVECLSD 714
Db 629 HPDKIEFTPTATLARESDLERAQKAVNALFTSSNQIGLKTDDVTDYHIDRVSNLVECLSD 688
Qy 715 ELYPNEKEMLDVAKEAKRLVQARNLLQDTGFNRING--ENGWGTGSGTGEVABGDVLPKD 772
Db 689 EFCLEKKEKSEKVKHAKRLSDERNLLQDPNFRGINRQLDRGWRGSTDTITQGGDDVPKE 748
Qy 773 RSLRLTSAREIDTETPTYLQOIDESELLKPYTRYKLGKFGYSSQDLBIKLIHRANOIV 832
Db 749 NVVTLGGTFD--ECYPTLYQKIDESKLAAYTRYQLRGYIEDSOOLBEIYLIRYNARHET 805
Qy 833 KNVPDNLPLDVLVUNSCGIDRCSQQYVDANLALENNGENGMSDSDSHAFSHIDTGEI 892
Db 806 VNVPGT--GSLWPLSAPSPIGKAHH-----SHHFLDIDVGCT 842
Qy 893 DLNENTGIWVVPFKIPTNGYATLGNLELVEEGLSGETLERAOQOQOQWODKMKRGAS 952
Db 843 DLNEDLGWVVFIPKIKTQDGHARLGNLEFLBEKPLVGEALAVKRAEKKWRDKREKLEWET 902
Qy 953 EKAYYAAKQADRLPADYQDQKLSGVEMSDMLAONLVQSIPYTYNDALPEIPGMNTYS 1012
Db 903 NIVYKEAKESVDALFVNSQYDRLOADTNIAITHAADKRVHSIRREAYLPPELSVIPGVNAAI 962
Qy 1013 FTELNRLOQANLVDLNAIPNGDFRNLGSLDWNATSDVNV-QQLSDTSVLVIPNNNSQV 1071
Db 963 FEELEGRIPTAFSLYDARNVKNKGDFNNGLSQVNVKGVHDVEEQNNHRSVLVPEWBAEV 1022
Qy 1072 SQOFTVQPNRYRLVRTARKEGVGDGYVIRDGANQTTETLPN-----ICDD 1118
Db 1023 SQEVVCPGCGYILRVATAYKEGYGCGVTIHEIENNTDELKESNCVEBEVYPPNVTVCND 1082
Qy 1119 DTGV-----LSADQTS----- 1129
Db 1083 YTATQBEYEGYTSRNRGYDGAYESNSSVPADYASAYBEKAYTDGRRDNPCESNRGYDY 1142
Qy 1130 -----YIKTVEFTPTSTQWIDMSBTEGVFNIESVELVLEEB 1167
Db 1143 TPLPAGYVTKLEYPFETDKWIEIGETGTFIVDSVELLMEB 1186

```

Search completed: June 21, 2004, 10:23:47
Job time : 417 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 10:10:20 ; Search time 85 Seconds
(without alignments)
4331.879 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYILDASSSTVS.....MSBTEGVNIESVELVLEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5903.5	97.7	1144	2	Q8KZL7	Q8KZL7 bacillus th
2	2750.5	45.5	1144	2	Q45745	Q45745 bacillus th
3	2545.5	42.1	1228	2	Q33T75	Q33T75 bacillus th
4	2539.5	42.0	1228	2	Q33NM5	Q33NM5 bacillus th
5	2433	40.3	1231	2	Q8KNY2	Q8KNY2 bacillus th
6	2056.5	34.0	1155	2	Q9F296	Q9F296 bacillus th
7	2050	33.9	1180	2	Q9S5V8	Q9S5V8 bacillus th
8	2048	33.9	1176	2	Q7MZ79	Q7MZ79 bacillus th
9	2045	33.8	1176	2	Q45736	Q45736 bacillus th
10	2034	33.7	1176	2	Q9RC30	Q9RC30 bacillus th
11	2032.5	33.6	1155	2	Q93T21	Q93T21 bacillus th
12	2016.5	33.4	1171	2	Q06894	Q06894 bacillus th
13	2003	33.1	1169	2	Q8GHE8	Q8GHE8 bacillus th
14	1997	33.0	1189	2	Q9L877	Q9L877 bacillus th
15	1989.5	32.9	1174	2	Q45749	Q45749 bacillus th
16	1953.5	32.3	1160	2	Q93TF9	Q93TF9 bacillus th

17	1948.5	32.2	1128	2	Q9PDC0	Q9fcd0 bacillus th
18	1939.5	32.1	1177	2	Q8GLY5	Q8g1y5 bacillus th
19	1937	32.0	1176	2	Q9S514	Q9s514 bacillus th
20	1933.5	32.0	1177	2	Q03743	Q03743 bacillus th
21	1933	32.0	1178	2	Q45768	Q45768 bacillus th
22	1931	31.9	1178	2	Q9R826	Q9r826 bacillus th
23	1930.5	31.9	1177	2	Q45735	Q45735 bacillus th
24	1925.5	31.9	1118	2	Q9AM83	Q9am83 bacillus th
25	1897.5	31.4	1118	2	Q9AM82	Q9am82 bacillus th
26	1836.5	30.4	1280	2	Q8VUK9	Q8vuk9 bacillus th
27	1780.5	29.5	1236	2	Q939T3	Q939t3 bacillus th
28	1746.5	28.9	1118	2	Q9AM81	Q9am81 bacillus th
29	1730.5	28.6	719	2	Q9F0P8	Q9f0p8 bacillus th
30	1727.5	28.6	719	2	Q85796	Q85796 bacillus th
31	1717.5	28.4	719	2	Q93NJ5	Q93nj5 bacillus th
32	1702.5	28.2	719	2	Q8KY61	Q8ky61 bacillus th
33	1702	28.2	1254	2	Q8VULO	Q8vul0 bacillus th
34	1577.5	26.1	1270	2	Q8VUL1	Q8vul1 bacillus th
35	1378.5	22.8	638	2	Q87654	Q87654 bacillus th
36	1278.5	21.2	1155	2	Q9AM80	Q9am80 bacillus th
37	1254	20.7	645	2	Q9S603	Q9s603 bacillus th
38	1245	20.6	652	2	Q9S6N9	Q9s6n9 bacillus th
39	1171	19.4	723	2	Q9S4B5	Q9s4b5 bacillus th
40	1094	18.1	489	2	Q8KNV1	Q8knv1 bacillus th
41	1069.5	17.7	533	2	Q7X3F6	Q7x3f6 bacillus th
42	1061	17.6	381	2	Q45740	Q45740 bacillus th
43	1052.5	17.4	558	2	Q8VW63	Q8vw63 bacillus th
44	1042.5	17.2	526	2	Q32308	Q32308 bacillus th
45	1041	17.2	620	2	Q45720	Q45720 bacillus th

ALIGNMENTS

RESULT 1

Q8KZL7	PRELIMINARY;	PRT; 1144 AA.
ID	Q8KZL7	
AC	Q8KZL7;	
DT	01-OCT-2002 (TRENBLrel. 22, Created)	
DT	01-OCT-2002 (TRENBLrel. 22, Last sequence update)	
DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)	
DE	Cry8 protein.	
GN	Cry8.	
OS	Bacillus thuringiensis (subsp. galleriae).	
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.	
OX	NCBI_TaxID=29338;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=SDS-502;	
RA	Asano S., Yamamoto T.;	
RT	"a novel cry8 gene highly toxic to Anomala cuprea.";	
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AB089299; BAC07226.1; -	
DR	GO; GO:0015070; F:toxin activity; IEA.	
DR	GO; GO:0006952; P:defense response; IEA.	
DR	InterPro; IPR001178; Endotoxin.	
DR	InterPro; IPR005638; endotoxin C.	
DR	InterPro; IPR008979; Gal_bind_like.	
DR	Pfam; PF00555; endotoxin_1.	
DR	Pfam; PF03944; endotoxin_C_1.	
DR	Pfam; PF03945; endotoxin_N_1.	
SQ	SEQUENCE 1144 AA; 128060 MW; 98F93070C49014AB CRC64;	

Query Match	97.7%;	Score 5903.5;	DB 2;	Length 1144;
Best Local Similarity	97.9%;	Pred. No. 0;		
Matches 1143;	Conservative	0;	Mismatches	1;
Indels	23;	Gaps	1;	
Qy	1	MSPNNQNEYILDASSSTVS	SDNSVRVPLANDQTTTLQNNYKDYLRMS	EGENPBLFGNP 60
Db	1	MSPNNQNEYILDASSSTVS	SDNSVRVPLANDQTTTLQNNYKDYLRMS	EGENPBLFGNP 60
Qy	61	ETFISSSTVQTGIVGQVLGALGVPFAGQIASFIS	PIVGQWPSSTVS	VWEMIMKQVED 120

Db	61	ETPISSTVQTGIGIVGOVLGALGVPPAGQTASFYSFIVGQLPSPSTSVWEMIMKQVED	120
Qy	121	LIDOKITDSVRKTAAGLQGLGDLGVYQKSLKQWLENRNDTRARSVVVYQYIALELDPV	180
Db	121	LIDOKITDSVRKTAAGLQGLGDLGVYQKSLKQWLENRNDTRARSVVVYQYIALELDPV	180
Qy	181	AKIPSPAISGOBVPLLSVYAAQANLHLLLRDASIFGAENGFTPGESTISTFYDQVTRTAQ	240
Db	181	AKIPSPAISGOBVPLLSVYAAQANLHLLLRDASIFGAENGFTPGESTISTFYDQVTRTAQ	240
Qy	241	YSDYCVWYNTGLDKLGTHAASHLKTHQFRRENTLLVLDLVALFPNDYDRTPIETTAQ	300
Db	241	YSDYCVWYNTGLDKLGTHAASHLKTHQFRRENTLLVLDLVALFPNDYDRTPIETTAQ	300
Qy	301	LTRREYVTDPIVFNKRETSGGFCRRWSLNSDISFSVESAVIRSHPHLPDILSEIEFYTTTRAG	360
Db	301	LTRREYVTDPIVFNKRETSGGFCRRWSLNSDISFSVESAVIRSHPHLPDILSEIEFYTTTRAG	360
Qy	361	LPLNNTLEYWVGHSIKYKXNTNASSALERNYGTITSNKIKYYDLANKDIFQVRSLGADL	420
Db	361	LPLNNTLEYWVGHSIKYKXNTNASSALERNYGTITSNKIKYYDLANKDIFQVRSLGADL	420
Qy	421	ANYAQQVYGPYASFTLLDKNTGSGVGGFTYSKPHITTMQVCTQNYNTIDEIIPENEPPLS	480
Db	421	ANYAQQVYGPYASFTLLDKNTGSGVGGFTYSKPHITTMQVCTQNYNTIDEIIPENEPPLS	480
Qy	481	RGYSHRLSHITSYSFSPKSNASSPARYGNLVPFAMTHRSADVTNTVYSDKITQIPVVKAAHTL	540
Db	481	RGYSHRLSHITSYSFSPKSNASSPARYGNLVPFAMTHRSADVTNTVYSDKITQIPVVKAAHTL	540
Qy	541	VSGTTVIKGPFGTGGNLLKRTSSGPLAYSVSVKSPLSQRYRARIYASTTNLRLPVTIS	600
Db	541	VSGTTVIKGPFGTGGNLLKRTSSGPLAYSVSVKSPLSQRYRARIYASTTNLRLPVTIS	600
Qy	601	GTRIYSINVKTMKGBDLPFTNTFDLATIGTAPTFPSYSDSLTVGADSPASGGEVYVDKF	660
Db	601	GTRIYSINVKTMKGBDLPFTNTFDLATIGTAPTFPSYSDSLTVGADSPASGGEVYVDKF	660
Qy	661	ELIIPVNATFPAEEDLDVAKKAVNGLFTSKDQALQTSVTDYQVNOAANLVCLSDSELYPNE	720
Db	661	ELIIPVNATFPAEEDLDVAKKAVNGLFTSKDQALQTSVTDYQVNOAANLVCLSDSELYPNE	697
Qy	721	KRMLWDVAKGAKRLVQARNLLQDTGPNRINGENGWTCSTGLEVABGDVLFKDRSLRLTSA	780
Db	698	KRMLWDVAKGAKRLVQARNLLQDTGPNRINGENGWTCSTGLEVABGDVLFKDRSLRLTSA	757
Qy	781	REIDTETPTLYQOQIDESLLKPYTRYKLGKFIGSSQDLEIKLIRHRANOIVKXNPONLL	840
Db	758	REIDTETPTLYQOQIDESLLKPYTRYKLGKFIGSSQDLEIKLIRHRANOIVKXNPONLL	817
Qy	841	PDVLPVNSCGGIDRCSQQYVDANLALNNNGENGMSSDHAQSPHIDTCGIEDLNENTGI	900
Db	818	PDVLPVNSCGGIDRCSQQYVDANLALNNNGENGMSSDHAQSPHIDTCGIEDLNENTGI	877
Qy	901	WVVPKIPITNGYATLGNLELVEBGLSGETLERAQQOQEQQWQDKMRKRGASEKAYYAAK	960
Db	878	WVVPKIPITNGYATLGNLELVEBGLSGETLERAQQOQEQQWQDKMRKRGASEKAYYAAK	937
Qy	961	QAIDRLPADYQDQKLSNGEVMSDMLAONLVQSIPYVYNDALPEIPGMNNTSFTELNRLL	1020
Db	938	QAIDRLPADYQDQKLSNGEVMSDMLAONLVQSIPYVYNDALPEIPGMNNTSFTELNRLL	997
Qy	1021	QOAWNLVDLRANAIPNGFPRCLSDMNATSDVNVQOOLSDTSVIATPKNNSOVSOQFTVQPN	1080
Db	998	QOAWNLVDLRANAIPNGFPRCLSDMNATSDVNVQOOLSDTSVIATPKNNSOVSOQFTVQPN	1057
Qy	1081	YRYVLRVTARKEGVDGTVIIRDOGANOTETLTFNICDDDTGVLSDAQTSVITKTVEFTPS	1140
Db	1058	YRYVLRVTARKEGVDGTVIIRDOGANOTETLTFNICDDDTGVLSDAQTSVITKTVEFTPS	1117
Qy	1141	TEQWVIDMSETEGVFNIESVELVEE	1167

Db 1118 TEQWIDMSTEGVFNIESVELVLEE 1144

RESULT 2

Q45745 PRELIMINARY; PRT; 1144 AA.

ID Q45745

AC Q45745

DT 01-NOV-1996 (TEMBLrel. 01, Created)

DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DB Delta-endotoxin (Pragmat).

DE CRYIX GENE.

OS Bacillus thuringiensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1428;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94085596; PubMed=8262221;

RA Shevelev A.B., Svarinsky M.A., Karasin A.I., Kogan Y.N.,

RA Chetukhina G.G., Stepanov V.M.;

RT "Primary structure of the cryX--the novel Delta-endotoxin-related

RT gene from *Bacillus thuringiensis* ssp. *Galleriae*.";

RL PDBS Lett. 336:79-82(1993).

DR EMBL; X75019; CAA52927.1; -.

DR HSP; P07130; IDLC.

DR GO: 0015070; P: toxin activity; IEA.

DR GO: 0006352; P: defense response; IEA.

DR InterPro: IPR001178; Endotoxin.

DR InterPro: IPR005638; endotoxin_C.

DR InterPro: IPR005639; endotoxin_N.

DR InterPro: IPR008979; Gal_bind_Like.

DR Pfam: PF00555; endotoxin; I.

DR Pfam: PF03944; endotoxin_C; 1.

DR Pfam: PF03945; endotoxin_N; 1.

FT NON TER 1

SQ SEQUENCE 1144 AA; 129399 MW; 7D28594A19C4B065 CRC64;

Query Match 45.5%; Score 2750.5; DB 2; Length 1144;

Best Local Similarity 48.7%; Pred. No. 2.8e-148;

Matches 562; Conservative 197; Mismatches 341; Indels 53; Gaps 18;

Qy 41 NYKDYLRMSGEMPELFGNPETETPISSTVGTGIGVQVLGALGVPPAGQIASPYSPIVG 100

Db 19 SYKDYLRMSGSDYIDSYINPENV--RTGLQTGIDIVAVVVGALGGPVGGILTGFLSTLFG 76

Qy 101 QLPSPSTVSVMEMKQVEDLDQKITDSVRKTALAGLOGLDGLDVYQSKLNWLENNR 160

Db 77 FLWPSNDQAVWZAFIEQMEELEIQRISSDQVVRTALDGLTGQIYNYQYLALKWESESRPN 136

Qy 161 DTPARSVVVQYTALELDLVAKTIPSPAISQ-----EVLPLSVYQAANLHLLLRDASI 215

Db 137 GVRA-NLVLRQRFILHALFVSSMPSFG-SGFGSQRFQQLLVYTAQANLHLLLRDREK 194

Qy 216 FGAEWGTFPGEISTFYDRQV--TRTAQVSDYCVKWYNTGLDKAGTNAASWLKYHQPREM 274

Db 195 YGARWGLRSEQIGNLYFENELQTRDYTNHCNVNNGLAGLGTSAESWLKYHQPREA 254

Qy 275 TLLVLDLVALFPNYDRTYPIETTAQLTRVYTDPIVFNRETSGP-----CRWSLNSDI 330

Db 255 TLMAMDIALFPYNTRRYPYIAVNPQITREVYTDPIGLGVPSSESLPPELRCLRMQETSAM 314

Qy 331 SPFEVESAVIRSPHLFDILSEIFPYYTTRAGLPLNNTLEYLVWVGHSI--KYKVTNASSAL 388

Db 315 TFSLENLAITSSPHLFDITNNLMYVYGSFVHLTN-QLIEGWNIGHSVTSLLASGPTTVL 373

Qy 389 ERNYGTSNKIKYDPLANKDIPQVRS---LGLADLANIYAQVGVYPVASTLLDKNXTGSG 445

Db 374 RRVVGSSTTS-IVNVFSPNDRDVQINRSHTGFGQN--APLFGITRAQF----- 420

Qy 446 SVGGFTYSKPHTNQVCTQYNTIDEP---PENEPISRGYSHRLSHITSYSF-----SK 497

Db 421 -YFGGTYVQTRNALTCQNYNSIDSLPDPNRPISRSYSHRLSHITSYLRVLTIDGI 479

DR GO: GO:0015070; F:toxin activity; IEA.
DR GO: GO:0006952; P:defense response; IEA.
DR InterPro: IPR001178; Endotoxin.
DR InterPro: IPR005638; endotoxin N.
DR InterPro: IPR005639; endotoxin N.
DR InterPro: IPR008979; Gal_bind_like.
DR Pfam: PF00555; endotoxin; 1.
DR Pfam: PF03944; endotoxin C; 1.
DR Pfam: PF03945; endotoxin N; 1.
SQ SEQUENCE 1155 AA; 130557 MW; 5D69E3E2F527749D CRC64;

Query Match 34.0%; Score 2056.5; DB 2; Length 1155;
Best Local Similarity 39.2%; Pred. No. 1.2e-108;
Matches 477; Conservative 184; Mismatches 389; Indels 167; Gaps 29;

QY 53 NPEL-----FGNPET-FISSSTVQTG---IGI---VGQVLGALGVPPAGQIASFYSP 97
DB 4 NPINECPYCNLSNPEVEVLGGERIEGTYPIDISLSITQPLLEFPVPGAGFVGLVDI 63
QY 98 IVGQLMPSTSVSWEMIMKQBEDLDQKITDSVRKTALAGLQGLDGLDVYQKSLKNLLE 157
DB 64 IWGIFGPQ---WDAFLVQIBOLINQRIEHPARNOAISRLBGLSNLYQIVAESPREMEA 119
QY 158 NRNDTRARSVVVTVIALELDFVAKIPSPAISQGVPLLSVYQAANLHLILLRDSIFG 217
DB 120 DPTNPALREEMRIQFNDMNSALTTPAIPFAVQYQVPLLSVYQAANLHLSVLDRDVSFG 179
QY 218 AEWGTPGCEISFYDQVTRTAQYSDYCVKWTNTGLDKLGTNAASWLKYHOPREMTLL 277
DB 180 QRWGDAATINSRYNDLRLIGNYTHAVRWNTGLRWGDPDSRDWRYNQFRELTLT 239
QY 278 VLDELVALPNDYTRTYPIETTAQLTREVYTDPIVFNRETSGFCRRWLSNDSIFSEVES 337
DB 240 VLDIIVSLFPNDYTRTYPIETTAQLTREVYTDPIVFNRETSGFCRRWLSNDSIFSEVES 337
QY 338 AVIRSPHLPDILSEIEFTTRAGLPLNTEYLEYVGHISIK-----YKTNASSALERNY 392
DB 290 S-IRSPHLMILNLSITTYT-----DAHRGEY--YWSGHQIMASPVGPSPEPTFPL---Y 338
QY 393 GTITSNKIKYDLANKDIPQVRLGADLANYYAQVY--GVPVASTPLTLDKNT---GSGS 446
DB 339 GTMGNAAPQORIVAGQGVYTLSTL---YRFPNIGINNQQLSVLDTGFAYGTSSN 395
QY 447 VGGFTYSKPHTTMQVCTQNYNTIDRIPPENE--PLSRGYSRHLSHITSY--SFSKNASPP 502
DB 396 LPSAVYKSGTV-----DSLDEIPQNNVPPRQGFHSLSHVSMFRSGFSNYSVI 447
QY 503 ARYGNLPVPAWTHRGADVTNTYSDKIQIPIVKAHTLVSGTIVKPGFTGCTNLIKETS 562
DB 448 IR---APMFSWTHRSABFNIIIPSSQITQIPLTKSTNLGSGTSVYKPGFTGCDLRRTS 504
QY 563 SGPLAYTSVSVKSPLSQRYARIRYASTTNLRLFTVTSIRIYSINVKNTMKNKGDDLTFN 622
DB 505 PQGISTLRNITAPLSQRYARIRYASTTNLRLFTVTSIRIYSINVKNTMKNKGDDLTFN 564
QY 623 TPDLATIGTAFPTFSNYSLSLTGADSPASGGEVYVDKFEIIPVNAITFEABEDLDVAKAV 682
DB 565 SPRTVGTPTFPNFGSSVFTLSAUVFSGNEVIDRIEFVPAEVTFEAYDLERAQKAV 624
QY 683 NGLFTSKKD-ALQTSVTDYQVNAANLVECLSDLEYDPEKRMNDVAKELVQARNLL 741
DB 625 NELFTSSNQIGLKTVDYTHIDQVSNLVECLSDLEYDPEKRMNDVAKELVQARNLL 684
QY 742 QDTGNFRING--ENGWGTSGTLEVAEGDVLFDKRSRLTSAREIDETVPTLYQOIDES 799
DB 685 QPNLRGRNQDRWRGSDTITIQGDDVFENYVTLTGTFD---ECVPTLYQOIDES 741
QY 800 LKPYTRYLKGFIGSSQBLEKLRRHRANQIVKRVNPNLLPDVLPVNSCGGIDRCBQQ 859
DB 742 KLKAVTRYQLRGYIBDSQBLELYLYRYNAKHETVNVPGT--GSLWPLSAPSPIGKCAHH- 798
QY 860 YVDANLALNNGENGMSSDSHAFSFHDTGTGIDLNENTGIWVPKIPPTNGCYATLGNLE 919

DB 799 -----SHFSLDIDVGCTDILNEELGVVVPFKIKTQDCHARLGNLE 838
QY 920 LVEEGFSGTTLERAQQOQEQWQDMARKRGASERKAYYAAKQAIIDLPLADYQDQKLNQSV 979
DB 839 FLEEKPLVGEALARVKRAEKKWRDKREKLEWETNIVYVKEKESVDALFINQSYDRLOADT 898
QY 980 EMSDMAAQNVLQOSIPYVYNDALPIPKGNYYTSPTELTNRLQOANWLYDLRNAIPNGDPR 1039
DB 899 NIAMIHAADKRVHSIRREAYLPSELSVIFGVNAAIFBELEGRISTAFSLYDARNVIRKNGDFN 958
QY 1040 NGLSDNATSDVNV--QLSDTSVLVTPNNNSQVSOQFTVQPNRYVYVLRVTARKEGVGDGY 1098
DB 959 NGLSCWNVKGRVDVEQNNHRSVLVPEAEASVSEVVCPCGRIILRTAYKEGYGEGC 1018
QY 1099 VIIRGANQTETLTN-----ICDDDTGV----- 1122
DB 1019 VTIHIEENTDELKFSNCVEBEVYNNVTTCNDYTATQBEYEGTYTSRNRGVDGAYESNS 1078
QY 1123 -LSADQTS-----YITKVEFTSTEQVWIDMSE 1150
DB 1079 SVPADYASAYEEKAYTDCRRDNPCHSNRNGYDTPPLPAGYVTKLEYPFETDKVWIEIGE 1138
QY 1151 TEGVENIESVELVLEE 1167
DB 1139 TEGTIVDSVELLLEE 1155

RESULT 7
Q955V8
ID Q955V8 PRELIMINARY; PRT; 1180 AA.
AC Q955V8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DB BtT84A1 crystal protein (Crystal protein CryIA).
CB BtT84A1.
OS Bacillus thuringiensis, and
OC Bacillus thuringiensis (subsp. sotto).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI TaxID=1428, 29340;
RN [1]
RN [2]
RN [3]
RN [4]
RC SPECIES=B.thuringiensis; STRAIN=T84A1;
RA Ogo M., Yamada S., Kobayashi Y., Shibata J., Nagamatsu Y.;
RT "Nucleotide Sequence of the Lepidoptera-toxic Protein Gene of Bacillus
thuringiensis subsp. dendrolimus T84A1."
RL J. Pac. Appl. Biol. Sci. Hiroshima Univ. 29:95-107(1990).
RN [3]
RN [4]
RC SPECIES=B.thuringiensis; STRAIN=T84A1;
RA Nagamatsu Y., Itai Y., Hatanaka C., Funatsu G., Hayashi K.;
RT "A Toxic Fragment from the Entomocidal Crystal Protein of Bacillus
thuringiensis."
RL Agric. Biol. Chem. 48:611-619(1984).
RN [4]
RC SPECIES=B.thuringiensis; STRAIN=T84A1;
RA Nagamatsu Y., Itai Y., Hatanaka C., Funatsu G., Hayashi K.;
RT "A Toxic Fragment from the Entomocidal Crystal Protein of Bacillus
thuringiensis."
RL Agric. Biol. Chem. 48:611-619(1984).
RN [4]
RC SPECIES=B.thuringiensis (subsp. sotto);
RA Zhong W.F., Cai F.Z., Yan W.Z., Zhang Z.X., Xiang Y.W.;
RT "A cryIA gene cloned from Bacillus thuringiensis serovar sotto";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.
DB EMBL; AB026261; BAA77213.1; --
DB EMBL; AF510713; AAM4305.1; --
DB PIR; A22798; A22798.
DB HSP; P02965; ICIV.
DB GO; GO:0015070; F:toxin activity; IEA.
DB GO; GO:0006952; P:defense response; IEA.
DB InterPro; IPR001178; Endotoxin.
DB InterPro; IPR005638; endotoxin N.
DB InterPro; IPR005639; endotoxin N.
DB InterPro; IPR008979; Gal_bind_like.
DB Pfam; PF00555; endotoxin; 1.
DB Pfam; PF03944; endotoxin C; 1.
DB Pfam; PF03945; endotoxin N; 1.
SQ SEQUENCE 1155 AA; 130557 MW; 5D69E3E2F527749D CRC64;

Db 857 AR LGNTLEFLEBEKPIVGEALLARVTRAEKWRDXREKLEWE TNIVYKEAKESVDALFYNSQY 916
 Qy 973 QKLSNGVMSDMLAAQNLVQSIPYVYNDALPEIPGMNYTSTFETLNRLOQANWLYDLRNA 1032
 Db 917 DRLOADNTIAMIAHAADKRVHSIREAYLPESLVIPGVNAAI FFELEEGRIPTAFSLYDARNV 976
 Qy 1033 IPNGDFRGLSDNNATSDNVN-COOLSDTSLVNT PNWNSQVSQOFTOPNTYRVLYVTARK 1091
 Db 977 IRKGDFFNGLSCWNVKGHDVBSQNNHRSVLVVPFEWEAEVSQBRVVCPCRGYILRVYAYK 1036
 Qy 1092 EGVGDGVYIIRDCANOTETILTFN-----ICDDTTCV----- 1122
 Db 1037 EGVGEGCVTHIEYNTDELKFSNCVEEYVFNNTVTCNDYATQBEYSGTYSRNRGYD 1096
 Qy 1123 -----LSADQTS-----YTKTVFTFSTEQ 1143
 Db 1097 GAYESNSSVPADYASAYEKKAYTDGRDNPESNRGVDYTPLPAGVYTKLEYFPETDK 1156
 Qy 1144 VTI DMSETEGRVFNIESVELVLEE 1167
 Db 1157 WTIEIGETEGTFIVDSVELLAMEE 1180

RESULT 8
 Q7WZT9 PRELIMINARY; PRT; 1176 AA.
 ID Q7WZT9 AC Q7WZT9 AC Q7WZT9 AC
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Delta-endotoxin.
 GN CRY1AA.
 OS Bacillus thuringiensis.
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1428;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ly30;
 RA Yao J., Zhang J., Song F., Chen Z., Li C., Huang D.;
 RL "A novel cry1Aa gene from Bt strain ly30."
 RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF384211; AAP80146.1; -
 SQ SEQUENCE 1176 AA; 133169 MW; 725C14BA9408B69B CRC64;

Query Match 33.9%; Score 2048; DB 2; Length 1176;
 Best Local Similarity 39.3%; Pred. No. 38e-108;
 Matches 480; Conservative 174; Mismatches 414; Indels 152; Gaps 26;

Qy 53 NPEL-----FGNDET-FISSSTVQTG---IGI---VGQVLGALGVPPAGQIAFSYSF 97
 Db 4 NFINNECIPYNCUSNPEVEVLGGERIETGYTPIDISLSLTFLLSFBVFGNGVFLGLVDI 63
 Qy 98 IVGQLPSSSTVSVMEMIMKQVELDIDKTTSDSRKTKALAGLQGLGDLVYQKSLKNWLE 157
 Db 64 IWGIFGSPQ---WDAPLVOIEQLINQRIEETEPARNOAISREGLSINLYQIYAESFREWA 119
 Qy 158 NRNDTRARSVVVYQYIALELDFVAKIPSPAISQBRVPLLSVYAAANLHLALLLDRDASIFG 217
 Db 120 DPTNPALREMRIQPDMDNSALTTAIPLFAVQNYQVPLLSVYQAAANLHLVLDRDVSFVG 179
 Qy 218 AEWGPTGELSTFYDQVTRTAQSYCYKWNVTGLDKLGTWAAASWLKTHQRRREMTLL 277
 Db 180 QRWGFDAAATINSRYNDLTRLIGNYTDYAVRWYNTGLERWGPDSRDWRVYNQPRRELTLT 239
 Qy 278 VLDIVALPNYDRTYPIETTAOLTEBVTDPVFNRETS-GGFCRWSLNSDISFSEVE 336
 Db 240 VLDIVALFNSYDSRRPIRTVSQLTRREIYNPVLNFDGSRFGMAORIBON----- 290
 Qy 337 SAVIRSPHLPDLISIEFTY-TRAGLPLNNTYLEYVWGHSHK-----YKNTNASSALER 390
 Db 291 ---IROPHLMDILNLSITDYDVRG-----PNTWSGHOITASPVGSGPEFAPLPG 339
 Qy 391 NYGTITSNKIKKYDLANKDIFQVRS-----LGADLANYYAQVYGVVPYASFTLLDKNT 442


```
Qy 621 FNTFDLATIGTAFPTFSNYSDSLTVGADSPASGGSEVYVDKZELIPVNAATFAEBEDLVAKK 680
Db 562 SSGSFRVTGFTTFFNFSGSVFTLSAHVFNSEVYIDRIEFVPAEVTFAEYDLERAQK 621
Qy 681 AVNGLFTSKKD-ALQTSVTDYQVNOAANLVECLSDLYPNKRMKMDAVKEAKRLVQARN 739
Db 622 AVNELFTSSNQIGLKTVDYDHYDQVSNLVECLSDLYPNKRMKMDAVKEAKRLVQARN 681
Qy 740 LLODTGPNRNG--ENGWTSSTGIEVAEGDVLKORSRLTSAREIDTETPTLYQQID 797
Db 682 LLODPNFRGINRQDRGWRGSDTITIQGGDDVFKNVYVLLGTFD---ECYPTLYQKID 738
Qy 798 ESLLKPYTRYKLGKFGISSQDLLEKLRHRANQIVKXVDPN-----LLPDVLPVNSCGGID 853
Db 739 ESKLKAYTRYQLRGYIEDSQDLLEYLYRYNAKHETVNVPGTSLWPLSAQSPIGKCGEPN 798
Qy 854 RCSEQQVVDANLALENGNG--NMSSDSHAFSPHIDTGEIDLNEMTGIWVVFPIPTNGY 912
Db 799 RCA--PHLEWNPDLDCSCRDGKCNHSHHPSLIDVGCCTDLNEDLGWVIFPKIKTQDGH 856
Qy 913 ATLGNLBELVEEGLSGTTLERAQQOQWQDNKARKGASEKAYYAAQALDRLFADYQD 972
Db 857 ARGNLEFLEBKPLVGEALARKVRAEKWRDKREKLEWETNIVYKEAKESVDALFVNSQY 916
Qy 973 QKLSNGVEMSDMLAAQNLVOSIPVYNDALPEIPCMNYTSFTLTNRLOQAWNLVDLNA 1032
Db 917 DOLQADNTNIAHIAADKRVHSIREAYLPBLSVPGVNAAI FEELEGRIFTAFSLYDARNV 976
Qy 1033 IPNGDFRNLGSDMNATSDVAV-QQLSDTVLVPNNMSQVSOQTPQPNRYVLRVARTK 1091
Db 977 IKNGDFNGLSCNVKGVHVEEQNORSVLVPEWEAEVSOEVVCPGEGYILRVYAK 1036
Qy 1092 EGVGDGVYIIRDCANOTETLTFN-----ICDD-----1118
Db 1037 EGYGEGCVTTHIEINNTDLKFSNVBEIYPPNVTVCNDYVNOEYGGAYTSRNRGN 1096
Qy 1119 DTGVLSDQTS-----1119
Db 1097 EAPSVAPDAVSUYEEKSYTDGRRNCPENFRNGRYDTPLPVGVYVTKSELYFPETDKVWIE 1156
Qy 1148 MSBTEGVNIESVELVLEE 1167
Db 1157 IGTEGTFTVDSVELLLEE 1176

RESULT 10
Q9RC30 PRELIMINARY; PRT: 1176 AA.
AC Q9RC30;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE 135 kDa insecticidal protein.
OS Bacillus thuringiensis (subsp. kurstaki).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID:29339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HD-1-02;
RA Hou B.K., Chen Z.H.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154676; AAD55382.1; -.
DR HSSP; P02965; 1CIY.
DR GO; GO:0015070; P:toxin activity; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR Pfam; PF008979; Gal_bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
```

```
SQ SEQUENCE 1176 AA; 133010 MW; FCBEO69D0B81D8C4 CRC64;
Query Match 33.78; Score 2034; DB 2; Length 1176;
Best local Similarity 39.28; Pred. No. 2.4e-107;
Matches 478; Conservative 173; Mismatches 417; Indels 152; Gaps 26;

Qy 53 NPDL-----FCNPET-FISSSTVQTG---IGI-----VGQVLGALGVFPFQAQIASFYSF 97
Db 4 NPHNECIPYCNLSNPEVEVLGGRIETGTPIDISLSLTQFLSEFVGAGFVLGLVDI 63
Qy 98 IVGQLWPSSTVSVMIMKQVEDLIDQKITDSVKTALAGLQGLDGLDVTQKSLNMLE 157
Db 64 IWGIFGFSQ---WDTFLVQIEQLINQIRIBEFARNAQISRLGLESLNYQIYAESPREVEA 119
Qy 158 NRNDTRARSVVVQYIALELDFVAKIPSAISGOEVLISVYQAANLHLLLRDASIFG 217
Db 120 DPTNPALREEMRIQFNDMSALTTAIPLLAQVQVPLLSVYQAANLHLSVLRDVSVFG 179
Qy 218 AEWGFTPEISTPYDRQVTRTAQYSDYCVKMYNTGLDKLKGNTAASMLKYHQFREMVTLL 277
Db 180 QRWGFDAAATINSRYNDLTRLIGNYTDYAVRYNTGLERWGPDSRDWRYNQFRELIT 239
Qy 278 VLDLVALFPNYDTRTYPIETTAQLTREYTDPIVFNRETS--GGFCRRWSLNSDISFSBVE 336
Db 240 VLDIVALFSNVDSRRYPRTVSQLTRBIYTNVPLETDFGSPRGMARIEQN-----290
Qy 337 SAVIRSHPLFDLISEIFYT-TRAGLPLNNTLEYLVVGHSHIK-----YKNTNASSALER 390
Db 291 ---IROPHLMDILNSITITVDVHRG-----PNYMSGHQITASPVGFSGPEFAPLFG 339
Qy 391 NYGHTITSNKLKYDLANKDIPQVRS-----LGADLANYYAQVYGVVPYASFLLDKNT 442
Db 340 NAGNAAPPVL--VSLTGLGIFRTLSSPLRYRILLCGSPNNQELFVLDTEFSASLTN 397
Qy 443 GSGSVGGFTYSKPHTTMQVCTONTNTIDEIPPENE--PLSRGYSHRLSHITSYSFSKNAS 500
Db 398 PS-----TIYRQGTV-----DSLDPVIPPQDNSVPPRAGFSHRLGHVMTLSQAAGAV 444
Qy 501 SPARYGNLPVPAWTHRADVTNVYSKTIQIPIVKAHTLVSGTIVIKGPGTGNILKR 560
Db 445 YTLR---APTFSQWHSABFNNIIPSSQITQIPLTKNSLNGSGTSVVRKPGFTGGDIIRR 501
Qy 561 TSSGPLAYTSVSVKSPLSQRYRARIYASTTNLRLFTVITSGTRIYSINVNNTMKGGDILT 620
Db 502 TSPQISTLRVNIITAPLSQRYRARIYASTTNLRFHTSIDGRPINQGNFSAITMSSGSNLQ 561
Qy 621 FNTFDLATIGTAFPTFSNYSDSLTVGADSPASGGSEVYVDKZELIPVNAATFAEBEDLVAKK 680
Db 562 SSGSFRVTGFTTFFNFSGSVFTLSAHVFNSEVYIDRIEFVPAEVTFAEYDLERAQK 621
Qy 681 AVNGLFTSKKD-ALQTSVTDYQVNOAANLVECLSDLYPNKRMKMDAVKEAKRLVQARN 739
Db 622 AVNELFTSSNQIGLKTVDYDHYDQVSNLVECLSDLYPNKRMKMDAVKEAKRLVQARN 681
Qy 740 LLODTGPNRNG--ENGWTSSTGIEVAEGDVLKORSRLTSAREIDTETPTLYQQID 797
Db 682 LLODPNFRGINRQDRGWRGSDTITIQGGDDVFKNVYVLLGTFD---ECYPTLYQKID 738
Qy 798 ESLLKPYTRYKLGKFGISSQDLLEKLRHRANQIVKXVDPN-----LLPDVLPVNSCGGID 853
Db 739 ESKLKAYTRYQLRGYIEDSQDLLEYLYRYNAKHETVNVPGTSLWPLSAQSPIGKCGEPN 798
Qy 854 RCSEQQVVDANLALENGNG--NMSSDSHAFSPHIDTGEIDLNEMTGIWVVFPIPTNGY 912
Db 799 RCA--PHLEWNPDLDCSCRDGKCNHSHHPSLIDVGCCTDLNEDLGWVIFPKIKTQDGH 856
Qy 913 ATLGNLBELVEEGLSGTTLERAQQOQWQDNKARKGASEKAYYAAQALDRLFADYQD 972
Db 857 ARGNLEFLEBKPLVGEALARKVRAEKWRDKREKLEWETNIVYKEAKESVDALFVNSQY 916
Qy 973 QKLSNGVEMSDMLAAQNLVOSIPVYNDALPEIPCMNYTSFTLTNRLOQAWNLVDLNA 1032
Db 917 DOLQADNTNIAHIAADKRVHSIREAYLPBLSVPGVNAAI FEELEGRIFTAFSLYDARNV 976
```

```
Qy 1033 IPNGDFRGLSDMNATSDVNV-QQLSDTSVLVIPNWSQVSOFTVQPNRYVILRVARTARK 1091
Db 977 IKNQDFNGLSCMNVGHVDVEBQNNQSRSLVVPWEAEVSEVRCVPCRGVILRVATYK 1036
Qy 1092 EGVGUGVVIIRDGANOTETLTPN-----ICDD----- 1118
Db 1037 EGYEGCVTHIEIRNTDDELKFSNCVBEIYSNNVTTCNDYTVNOBEYGGAYTSRNRGYN 1096
Qy 1119 DTGVLSDADOTS-----YITKTVETFPSTEQWID 1147
Db 1097 EAPSPADAYSVYEKSVYDGRRENPCFNRGRYDTPLPVGYVTKELYEPETDKWIB 1156
Qy 1148 MSETGVFNIESVELVLEE 1167
Db 1157 IGETGTFIVDSVELLME 1176

RESULT 11
Q93T21 PRELIMINARY; PRT; 1155 AA.
AC Q93T21;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Crystal protein CryIab16.
GN CryIab16.
OS Bacillus thuringiensis (subsp. israelensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1430;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=407;
RA Yu J., Tan L., Wu D., Pang Y.;
RT "Molecular characterization of a silent gene encoding a 130-kilodalton
RT crystal protein from Bacillus thuringiensis subsp. israelensis.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375608; AAK55546.1; -.
DR F01; A29125; A29125.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin C; 1.
DR Pfam; PF03945; endotoxin N; 1.
SQ SEQUENCE 1155 AA; 130747 MW; 7F0C98E0100C7698 CRC64;

Query Match 33.6%; Score 2032.5; DB 2; Length 1155;
Best Local Similarity 38.9%; Pred. No. 2.9e-107;
Matches 473; Conservative 185; Mismatches 392; Indels 167; Gaps 29;

Qy 53 NPGL-----PGNPET-FTSSSTVQVG---IGI---VGQVILGALGVPPAGQIASFVSP 97
Db 4 NPNTNECIPYCNLSNPVEVLGGERIETGYTIDISLQTPLLSEFPVPGAGFVLGLVDI 63
Qy 98 IVGQLWPSSTSVSMKMKQVBDLIDOKITDSVRKTALAGLQGLGDLVDYQSLKNWLE 157
Db 64 IWGIFGPGSQ----WDAFLVQIEQLNQRIEERFARNQASRLGLESLNLYQIYAESPREWEA 119
Qy 158 NRNDTRASVVVTVQYIALELDFVAKIPSPAISQGVPLSVYAQAANLHLRLLRDASIFG 217
Db 120 DPTNPALREEMRIQPNDMNSALTATPLFAVQNYRVPVLLSVYVQAVNHLNLSVLVDLVFG 179
Qy 218 AEWGTFPGHISTFYDQVTRTAQYSDYCVKWTNTGLDKLKTNAASWLKYHQPREMTLL 277
Db 180 QRWGFDAATINSRYNDLRLIGNYTHAVRWYNTGLERVMGPDSDRDYRINQFRELTLT 239
Qy 278 VLDLVALPNTYTRTYPIETTAQLREVVTDPVFNRETSGGFCRWSLNSDISPSEVES 337
Db 240 VLDIVSLFPNYSRTYPIRTVSQLTREIYNPLENFD--GSF---RGSAGQ-----IEG 289
```

```
Qy 338 AVIRSPHLFDILSBIFFYTTRAGLPLNNTYLEYVWGHSHK-----YKNTNASSALERNY 392
Db 290 S-IRSPHLMDILNSITTYT-----DAHRGEY---YMSGHQIMASPVGFSGPEFTFPL---Y 338
Qy 393 GTIISNKIKYDILANKOIPQRSIGADLANYYAQVY--GVVPASFTLLDKMT-----GSGS 446
Db 339 GTMGNAAPQORIVAQLOGGVYRTLSSTL---YRPFNIGINNQQLSVLDTGTEFAVGTSTN 395
Qy 447 VGGFTYKPHHTMQVCTQNTYNTIDBIPPENB--PLSRGYSHRLSHITSY--SPSKNASSP 502
Db 396 LPSAVYRKSTV-----DSLDEIPQNNVPPRQGFSHRLSHVSMFRSGFSNVS61 447
Qy 503 ARYGNLVPFAWTHRSADVNTVYSDKITQIPVKAHTLVSGTIVTKGPGTCGNTILKRTS 562
Db 448 IR--APMFSWIHRSABFNIIIPSSQITQIPLTKSTNLGSGTSVVKGPGFGTGLLRRTS 504
Qy 563 SGPLAYTSVSVKSPLSQRYRARIYASTTNLRLFTVITGTRIYSINVNKTNKGDDLTEN 622
Db 505 PQQISTLAVNTAPLSQRYRARIYASTTNLQFHTSIDGRPINQGNPFSATMSGNSLQSG 564
Qy 623 TFDLATIGTFTFSPNSDSLTVGADSPASGGEVYVDKPELIPVNATPFAEBDLDVAKAV 682
Db 565 SFRTVGFTTPNFNSGSSVFTLSAHVFNSGNEVYIDRIEFVPAEVTFRABYDLERAQKAV 624
Qy 683 NGLFTSKD-ALQTSVTDYQVNOAANLYECLSDLEYLPHKEKMLWDVAKELVQARNLL 741
Db 625 NELFTSSNQIGLKTVDVTDYHIDQVSNLYECLSDLEYLPHKEKMLWDVAKELVQARNLL 684
Qy 742 QDTGFNRING--ENGWGTSGTGEVAGDVLPKDRSLRLTSAREIDTETPTLYLQOQIDES 799
Db 685 QDPNFRGINQLDRGWRGSTDITIOGGDDVPKENTVTLTGTFD---ECYPTLYLQKIDES 741
Qy 800 LLKPYTRYKLKGFIGSSQDLKILIRHRANQIVKQVNDNLDPVLVFNWSCGIDRCSEBQ 859
Db 742 KLKATRYQLAGYIBDSQDLRIYLIRYNAGHETVNVPGT--GSLWPLSAPSPIGKCAHH- 798
Qy 860 YVDANLALENNGENGWSDDSHAPSFHIDTGEIDLANENTGIWVVPKIPPTNGVATLGNLE 919
Db 799 -----SHFSLDIDVCGCTDLNEDLGWVVIPIKIKTODGCHARLGNLE 838
Qy 920 LVEEGPLSGETLERAQOQBOOWDMARKGRGASERKAYVAAKQAIIDRLFADYDQKLNISGV 979
Db 839 FLEEKPLVGEALARVKRAEKWRDKREKLEWTNIVYKEAKESVDVAVFVNSQYDLQADT 898
Qy 980 EMSDMIAAQNVLQSIPTVYVNDALPBIPGNNTYFTFELTNRLQOQANLYDLRNAIPNGDPR 1039
Db 899 NIAMIHAADKRVHSIRBAYLPBLSVPIGVNAAIPELEGRITAPSLDYARNVIRKNGDFN 958
Qy 1040 NGLSDWNATSDVNV-QQLSDTSVLVIPNWSQVSOFTVQPNRYVILRVARTARKBGVDGY 1098
Db 959 NGLSCMNVKGRHVDVEEQNNHRSVLVVPWEAEVSEVRCVPCRGVILRVATYKBEYBGC 1018
Qy 1099 VIIRGAMNOTETLTPN-----ICDDDTGV----- 1122
Db 1019 VTIHIEENTAELEKSNCEBEVYPNNTVTCNDYATATQEBYEGTYSRNRGYDGAYESNS 1078
Qy 1123 -LSADOTS-----YITKTVETFPSTEQWIDMSE 1150
Db 1079 SVPADYASAYEAKYATDGRRONPCBSNRGVDYTPPLPAGYVTRLEYFPETDKWIBIGE 1138
Qy 1151 TEGVFNIESVELVLEE 1167
Db 1139 TEGTFIVDSVELLME 1155
```

RESULT 12

```
Q06894 PRELIMINARY; PRT; 1171 AA.
ID 006894
AC 006894;
DT 01-JUL-1997 (TremBLrel. 04, Created)
DT 01-JUL-1997 (TremBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
```


305 NNITFTDFSVGRN-----FYMGHHRVSSLIIGGNNITSPYGRZANQBPFRSFTENG 359
410 IFQVRSIGADLANYYAQQVGVYAFLLDKQGTSGVGGFTYKPHFTTQVCTQNTI 469
360 VP--RILSNFTLRLLQWPAPPPNLGVBGVFTPTNSFTYGRGTV-----DSL 409
470 DEIPPENE--PLSRGYSHRLSHITSYSFKNASSPARYNLVPFAMTHRSADVNTVYSD 527
410 TELPPEDNSVPREGYSHRLCHAI---FVQRSGTPTLTGCV-VPSWTHRSATLTNTIDPE 465
528 KLTQIPVKAHTLVSTTVTKGPGTGNILKRTSSGFLAYTSVSKSPLSQRYRARI 587
466 RINQIPLVGRFRVWGTSVTGPGTGGDILRNTFGDVFSLQWNINSPIQRYLFRY 525
588 ASTNLRLFTV-----ISGTRIYSINVKTKNGKDDLTFTNTFDLTATGATPNSYSD 640
526 ASRDARVIVLGAATGCVGGQVSNMPLQKTEIGENLTSRTFTYDFSNPFSFRANPD 585
641 SLTV-----GADSPAGSGEVYVDKPELIPVNAATFEAREDLVAKKAVNGLPTS-KKD-AL 693
586 IIGISERPLFGAGSISS-GELYDKIEILADATFEAREDLVAKKAVNGLPTS-KKD-AL 693
694 QTSVTDYQVNOANLVCLSDLELYPNEKMLMDVAKKRLVOARNLLQDTGFNRIN--G 751
645 KTDVTDYHIDQVSNLVKCLSDLELYPNEKMLMDVAKKRLVOARNLLQDTGFNRIN--G 751
752 ENGTGSGTGIEVAGDVLKORSRLTSABRITETPTVLYQIDESLLKPTRYKLKG 811
705 DRGWRGSDTITQGGDDVFKENVTLPGT--VD-ECYTYLYQIDESLLKPTRYKLKG 761
812 FTGSSQDLKILIRHRANOIVKNVDPN-----LIPDLVFNVSCGIDRCSEQQVVDANLAL 867
762 YIEDSQDLKILIRHRANOIVKNVDPN-----LIPDLVFNVSCGIDRCSEQQVVDANLAL 867
868 ENNGENG-NWSSDSHAFSHIHTGETIDNENTGIWVFKIPTNGYATLGNLELVEEGPL 926
820 DCSRDGKCAHSHHFTLIDVCGTDLNEDLGWVIFKIQDGHARGNLNLELVEEGPL 879
927 SGETLERAOQOQODKMKRKGSEKAYAAQKQIDRLPADYQDQKNSGVHSDMLA 986
880 LGEALARVKAERKWRDKRKLQLETNIVYKAKESVDALFVNSQYDRLQVDTNAMIHA 939
987 AQNLVQSIPYVNDALPRIPKANYTSFTLTVRLAQAMNLYDLRANAIKNGDFRNLSDWN 1046
940 ADKRVHRIEALPELSVIFGVNAAIFEELEGRITAYSLYDARVINKNGDFRNLSDWN 999
1047 ATSDVNV--QQLSDTSVLVFNWNSQVQFTVQPNRYVLRVTKRGVGDGWIIRDCG 1105
1000 VKGHVDVEQNNHRSVLVPEWAEVQSVQVCPGCGVILRTVAYKEGVGSCVTIHEIE 1059
1106 NOTETLTN-----ICDDDTGV-----LSADOT 1128
1060 DNTDELKFSNVEEVEVPNTVTCTNNTYGTQBEYEGTYSRRNQGYDEAFGNPNVPADYA 1119
1129 S-----YITKVEFTPTSTROVWIDMSETGEVNI 1157
1120 SVYBEKSYDGRRENPCESNRYGDTPLPAGYVTKDLETFPTDKWIRIEGETGTPIV 1179
1158 ESVELVLEER 1167
1180 DSVELVLEER 1189

RESULT 15
Q45749
ID Q45749 PRELIMINARY; PRT; 1174 AA.
AC Q45749;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Crystal protein (Crystal delta-endotoxin).
GN CRY1PB.
OS Bacillus thuringiensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BTS00349A;
RA Lambert B.;
RT "NO INFORMATION";
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=B-Pr-88;
RA Li C., Zhang J., Huang D., Li G.;
RT "A crystal endotoxin from Bacillus thuringiensis strain B-Pr-88";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z22512; CAA80235.1; -;
DR EMBL; AF336114; AAO13295.1; -;
DR PIR; S32649; S32649.
DR HSP; P02965; ICIY.
DR GO; GO:001070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin C; 1.
DR Pfam; PF03945; endotoxin N; 1.
SQ SEQUENCE 1174 AA; 133350 MW; 8C7F122F9446F15C CRC64;

Query Match

32.9%; Score 1989.5; DB 2; Length 1174;

Best Local Similarity 38.0%; Pred. No. 8.4e-105;

Matches 458; Conservative 191; Mismatches 413; Indels 143; Gaps 25;

59 NPETFISSSTVQTG-----IGIVQVLGALGVPPAGQIASFYSTVQGLMPSTSVW 111
17 NPEVEILSERSTGRPLDLSLTFLLSEF-VPGVGVAFLDLMGFTTPE-----W 71
112 ENIMKQVEDLIDQKITDSVRKTALAGLQGLDGLVYQKSLKNLNRNDTRARSVVVQ 171
72 SLFLQLQIEQLQRIETLERNRAITTLGLADSYEVYLEALREWEENPNNAQLREDVIR 131
172 YLALSLDFVAKIPSPAISQVFPILLSVYQAANLHLLLRDASIFGAEMGFTTPEISTFY 231
132 FANTDADALITAINNFTLSFBIPLLSVYVQAANLHLLLRDASIFGAEMGFTTPEISTFY 191
232 DRQVTRTAQSYDVCVKNYNTGLDKLGTNAASWLKXHFQRRMTLAVLVALFPNYDTR 291
192 NRLNLIHRYTEHCLDTTQGLNLRGNTTROWSRPNQPRRLTLTLVLDIVALLFPNYDAR 251
292 TYPIETTAQLTREVTYDPIVFNRETSGGFCRWLSNLSDISPSESVSAVIRSPHLPDILSE 351
252 AYPITQSQTREIYVTSVIEDSPVSA-----NIPNGFNRAEFGV-RPHLADPMNS 302
352 IEFVTRAGLPLANNTEYLEVWGHISIKYKNTVASSALERNYGTIT-----SNKIY 402
303 L-FVTA-----ETVRSQTVMGGLVSSRNAGNPINFPYIGIPNGGAIWADHPRPP 355
403 YDLANKOIFQVRSIGADLANYYAQQVGVYAFLLDKQGTSGVGGFTYKPHFTTQVCTQNTI 469
356 YRTLSDPVF-VRG-GFGNPHVTLGRGVAF-----QQTGTN-----HTRTFRN 396
463 TONYNTIDEIPPENEPLS--RGYSHRLSHITSYSFKNASSPARYNLVPFAMTHRSADV 520
397 SGTIDSLDEIPPDNSGAPNDYSHVLNHTVFRWPGELAGSDSN-RAPMFWTHRSADR 455
521 TNYVYSDKITQIPVKAHTLVSTTVTKGPGTGNILKRTSSGFLAYTSVSKSPLSQRYRARI 587
456 TNIINPNIIITQIPAKYKAHNLHSGSTVVRGPGTGGDILRNTFGDVFSLQWNINSPIQRYLFRY 525
581 YRARIYASTTNLRLFTVTSIGTRIYSINVKTKNGKDDLTFTNTFDLTATGATPNSYSD 640
516 YRVRIYASTTDLQFPTTRINGTSVNOGNFQRTMRNGNLESGNFRTAGTFTSPFSFSAQS 575

Qy	641	SLTVGADSPAGBHVYVDKTELIPVNAITPEAREEDLVAKAVNGLFTSKD-ALQTSVTD	639
Db	576	TFTLGTQAF--SNQEVYIDRIEFVPAVTFPBAESDLERAQKAVNALFTSTSLGLKTNVT	634
Qy	700	YQVQAAANLVECLSDLEYLPNEKMLVDVAKELVQARNLLQDTGFNRLN--GENGWTG	757
Db	635	YHIDQVSNLVACLSDFCLDEKLSSEKVKHAKELSDKRNLLQDPNFRGNNRQPDHGWRG	694
Qy	758	STGIEVABGDVLPKORSRLRTSAREIDTETVPTLYIQOIDEGLKPKVTRYKLKGFIOSSQ	817
Db	695	STDITIOGGDDVPKENTVTLPGTFD---ECVPTLYQKIDESLKAAYTRYQLRGYIEDSQ	751
Qy	818	DLEIKLIRHRANQIVKNVP--DNLLPDVL--PVNSCGGIDRCSEQQVVDNLANLENNGEN	873
Db	752	DLEIYLLIRYNSKHEIVNVPGTGSLWPLSVENQIGPCGEPNCA--PHLEWNPDLHCSCRD	809
Qy	874	GNMS--SDSHAFSPHIDTGEIDLMENTGIWVVPKIPITNGVATYATLGNLVBEGPLSGETLE	932
Db	810	GERCVHSHHFSLDIDVGCTDLNEDLGVMILFKIKTQDGHARLGNLSFLESEPLGREALA	869
Qy	933	RACQOEOQODKWARKGASKEYAAKQALDRLFADYQOQKLSGVEMSDMLAAQHLVQ	992
Db	870	RVKRAEKKWDKREKLQLETNIVYKEAKESVDALFVNSQYDRLQADTNIAHRAADRKH	929
Qy	993	SIPVYNDALPEIPGMNYTSFTELTNLQAAWNLIDLNAIPNGDFRNLSDNNATSDVN	1052
Db	930	RIREALPELSVTPGVNAAIPFEELEGHIFATSYLDARNVTKGNFNGLLCNWVKGHVD	989
Qy	1053	V-QQLSDTSVLVIPNNNSQVSOQPTVQPNRYVLRVTARKEGVDGYIIRDCANQOTETL	1111
Db	990	VBEQNNHRSVLVIPWEAEVRSQKVRVCPGRGYILLRVTAKEGYEGGCVTIHEIEDNTDEL	1049
Qy	1112	TF-----	1113
Db	1050	KFSNCVBEGYPNNVTVTCHETMNOVGECTDACNVRNURGEDAYGHNPSTPVHYTTPYEE	1109
Qy	1114	-----NICDDDTGVLSAD--QTSYITIKVBEFTPTBQVWIDMSETBGFVNIESVEL	1162
Db	1110	ETYTDREPRNPCEANKGVNVTPLPVGYVTKYLEYFPETDTVMIEIGETBGTFLVDSVEL	1169
Qy	1163	VLREE 1167	
Db	1170	LLMEE 1174	

Search completed: June 21, 2004, 10:12:34
Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:35:05 ; Search time 28 Seconds
(without alignments)
4009.125 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYILDASSTSV.....MSETEGVNIESVELVLEE 1167

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues 283366
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3733.5	61.8	1160	2 I40589	parasporal crystal
2	3123	51.7	1157	1 S49247	parasporal crystal
3	2750.5	45.5	1154	2 S39536	parasporal crystal
4	2547.5	42.1	1228	2 S00873	parasporal crystal
5	2458.5	40.7	1138	2 A48944	parasporal crystal
6	2191	36.3	1156	2 S19306	parasporal crystal
7	2154.5	35.6	1174	2 A42459	parasporal crystal
8	2093.5	34.6	1166	2 S32645	parasporal crystal
9	2064.5	34.2	1155	2 A26513	parasporal crystal
10	2061.5	34.1	1155	2 JD0002	parasporal crystal
11	2052.5	34.0	1155	2 S02134	parasporal crystal
12	2048	33.9	1176	2 JT0241	parasporal crystal
13	2045	33.8	1176	2 JC2219	parasporal crystal
14	2044.5	33.8	1181	2 A41052	parasporal crystal
15	2043	33.8	1156	2 A29125	parasporal crystal
16	2038	33.7	1176	2 S02215	parasporal crystal
17	2037	33.7	1176	2 A22617	parasporal crystal
18	2030.5	33.6	1155	2 I39838	parasporal crystal
19	2024.5	33.5	1171	2 I40572	parasporal crystal
20	2010.5	33.3	1171	2 A37829	parasporal crystal
21	1989.5	32.9	1174	2 S32649	parasporal crystal
22	1984	32.8	1189	2 S00944	parasporal crystal
23	1953.5	32.3	1160	2 S32647	parasporal crystal
24	1944	32.2	1165	2 S11446	parasporal crystal
25	1939	32.1	1178	1 USBXKH	parasporal crystal
26	1932	32.0	1156	2 A29838	parasporal crystal
27	1930.5	31.9	1177	2 A49785	parasporal crystal
28	1923.5	31.8	1176	2 A48970	parasporal crystal
29	1917	31.7	1172	2 S32689	parasporal crystal

ALIGNMENTS

RESULT 1

I40589

parasporal crystal protein cry8Cal - Bacillus thuringiensis

N;Alternate names: parasporal crystal protein cryIII

C;Species: Bacillus thuringiensis

C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 01-Dec-2000

C;Accession: I40589

R;Sato, R.; Takeuchi, K.; Ogiwara, K.; Minami, M.; Kaji, Y.; Suzuki, N.; Hori, H.; Asan

Curr. Microbiol. 28, 15-19, 1994

A;Title: Cloning, heterologous expression, and localization of a novel crystal protein

A;Reference number: I40589; MUID:94100786; PMID:7764305

A;Accession: I40589

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1160 <RES>

A;Cross-references: EMBL:U04366; NID:G532523; PIDN:AAA21119.1; PID:G532524

C;Superfamily: parasporal crystal protein

C;Keywords: delta-endotoxin

Query Match 61.8%; Score 3733.5; DB 2; Length 1160;
Best Local Similarity 64.2%; Pred. No. 1.6e-208;
Matches 767; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

QY	1	MSPNNQNEYILDASSTSVSDNSRYPLANDQTTLQNNYKDYLRMSGEENPELPGNP	60
DB	1	MSPNNQNEYILDASSTSVSDNSIRYPLANDQTTLQNNYKDYLRMSTNAELSRNP	60
QY	61	ETPISSTVQTCIGIGVGLGALGVPPAGQIASFYFIVGQLPSPSTVSVMIMKQVE	119
DB	61	GTPIAQDAVGTGIDIVSTIISGLGIPVLGEVFSILGSLGLLMPNNENWQIPMKRVE	120
QY	120	DLIDQKITTSVRKTLAAGLGDLGVQKSLKQMLNENDTRARSVVVQYIALELDF	179
DB	121	ELIDQKILDSVRGRATADLANSRIAVFYQNALEDKRNPHSTSAALVKRFGNAAIL	180
QY	180	VAKIPSPAISGQRPVLLSYVAQANLHLLLRDASIFGAEWGFTPGIISTPTDQVTTA	239
DB	181	RTNMGSPSQNTYETPLPTVAQAASHLLVMDRVQIYKQWGPQNDIDLFYKQVSYTA	240
QY	240	OYSDYCVKQVNTGLDKLTNAASWLKHOPREMTLLVLDLVALFNYDTRTYPIETTA	299
DB	241	RYSDHCQVMYAGNKLGRGAKQWVDYRFRMNMVLDLVALFNYDARIIPLETNA	300
QY	300	QLTRVYVTDPIVFNRETSGGFCRRWSLN-----SDI-----SPSEVSASIRSPHFDIL	349
DB	301	ELTRVYVTDPIV-----GSYVTGQSTLSWYDMIPAAALPSPFTLEN-LLRKDPFFTL	352
QY	350	SEIEFTYTRAGLPLNNT-ELYEYVWGHISIKYKNTNASSALERNYGTITTSNKIKYDILANK	408
DB	353	QEIRMVTS---PRQMGTYEYNYWGGQRLTSLYIYGSSP--NKYSGLVAGAEIDIPVQGN	407
QY	409	DIFQVRSGLADLANYYAQVGVYPASFTLIDKNTGSGSGVGGFTYSKPHTTMQVCTQNYNT	468

Db 995 IPGIVEIYTELSDRLQQAQSYLYTSRNVAVQKGFNSGLDSWNTTMDASVQDQGNMPLVL 1054
 Qy 1065 PNNWSQVQFTQPNRYRYVLTARKEGVGDGVIIIRDGANOTETITFNICDDDTGVL 1124
 Db 1055 SHWDAQVQQLRVNPNCKYVLTARVGGGQGVYITRDGAHQETITFNAACDYVNGTY 1114
 Qy 1125 ADQTSYITKTVETPPTPEQWIDMSETEGPNVRSVELVLEE 1167
 Db 1115 VNDNSYITEVWPYPTKMMVSESEGSFYIDISFIEFIOE 1157

RESULT 3
 S39536
 parasporal crystal protein cry9Ba1 - Bacillus thuringiensis
 N;Alternate names: delta-endotoxin-related protein; parasporal crystal protein cryX
 C;Species: Bacillus thuringiensis
 C;Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 01-Dec-2000
 C;Accession: S39536
 R;Shevelev, A.B.; Svarinsky, M.A.; Karasin, A.I.; Kogan, Y.N.; Chestukhina, G.G.; Stepan
 FEBS Lett. 336, 79-82, 1993
 A;Title: Primary structure of the cryX, the novel delta-endotoxin-related gene from Baci
 A;Reference number: S39536; MUID:94085596; PMID:8262221
 A;Accession: S39536
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1154 <SHE>
 A;Cross-references: EMBL:X75019
 C;Superfamily: parasporal crystal protein
 C;Keywords: delta-endotoxin

Query Match 45.5%; Score 2750.5; DB 2; Length 1154;
 Best Local Similarity 48.7%; Pred. No. 1.9e-151;
 Matches 562; Conservative 197; Mismatches 341; Indels 53; Gaps 18;

Qy 41 NYKDYLRMSGEENDELFCNPEFTTSSSTVGTGIGVQVGLGALVPAGQIASFTSPVIG 100
 Db 29 SYKDYLRMSGEEDYDINFGNV--RTGLQGTIDIVAVVGLGALGPPVGGILTGFLSTLFG 86
 Qy 101 QLWPSSTVSVNEMIMKQVEDLIDOKITDSVRKTLALAGLGLDGLDVYQKSLKWLNRN 160
 Db 87 FLWPSNDQANWEAFIEQHEHLEIQRISDVVRLTALDGLTQNYNYLYALKWEERPN 146
 Qy 161 DTRARSVVVQYIALELDFAKIPSPALSGQ-----EVLPLSVYQAQANLHLLLRASI 215
 Db 147 GVRA-NLVQLRFTLHALFVSSMPSFG-SGPGSQRFQALLVYVYQAQANLHLLLRDAEK 204
 Qy 216 FGAEWGTPGEISFTYDROV--TRTAQYSDYCVKNTVGLDLKCTNAASMLKXHFQREEM 274
 Db 205 YGARWGLRESQIGNLYFNEIQTTRDYTNHCVNAYNNGLAGLGTSAESMLKXHFQREEA 264
 Qy 275 TLLVLDLVALPPNYDTRTYPIETTAQLTREYVTDPIVFNRETSGF-----CRRWLSNDSI 330
 Db 265 TLMAMDILALFPYNTTRYPIAVNPQLTREYVTDPLGVPSSESLFPELRLCLRINQWETSAM 324
 Qy 331 SPFSVESAVIRSPHLFDILSBIHFYTRAGLPLANNTEYLEYVGHSI--KYKNTNASSAL 388
 Db 325 TFSNLENAIISPHLEFDITINNLMYTGFSVHLTN-QLIEGWIGHSVTSLSLGSPTTVL 383
 Qy 389 ERNYGTTTSNKIKYIDLANKDIFQVRS---IGADLANVYQVYGVYPVASFLLDRNTGSG 445
 Db 384 RRYNGSTTS-IVNYFSFNDRDVYQINTSRHTGLGFQN--APLFGITRAQF----- 430
 Qy 446 SVGGFTYSKPHTTMQVCTQNTYNTIDEIP--PENPLSRGYSRHLSHITSYSF-----SK 497
 Db 431 -YPGTTYSVTORNALTCQNTNSIDELSPDNPEFSKYSRHLSHITSYLRVLTIDGI 489
 Qy 498 NASSPARYGNLPVFAWTHRSADVNTVYSDKITQIPVVKATRLVSGITTVIKGPGFTGNI 557
 Db 490 NIYS-----GNLPTVWTHRDVLTNTIADRTQLPLVKSFEIPAGITTVVRGPGFTGDI 545
 Qy 558 LKRTSSGPLATYSVSKPPLSQRTARIRYASTTNLRLFTVITSGTRIYSINNTKMKGD 617
 Db 546 LRRTGVTGFTGIRVTRTAPLQRYIRFRFPASITNLFLGIRVGRQVNYFQFGRTMNRGD 605

Qy 618 DLTFNTFDLATICTVAFTPFSNYSDSLTVGADSPFASGGVYVDKFKELIPVNTATFEAEEDLV 677
 Db 606 ELRYBSFATREFTTDFNFRQFQELISVFANAFSAGQEVYFDRIBIIPVNPAREAKEDLEA 665
 Qy 678 AKAVNGLPFSKDALQTSYTDYVNOAANLVECLSDLELYENKRMMLWDVAKRAKLVOA 737
 Db 666 AKKAVASLFRTRDGLQVNVKYQVQQAANLVSCLSDEQYGVYDKKMLLEAVRAAKRLSRE 725
 Qy 738 RNLQDTGFNRKNG--ENGMWTGSTGIEVAGDVLFDORSLSLTSAREIDTSTPTLYLQQ 795
 Db 726 RNLQDPPDFNTINSTENGWKAANGVTISEGGPYKGRALQLASAR-----ENYPTIYQK 781
 Qy 796 IDSELKPYTRYKLGKFGIGSSQDLLEIKLIRHRANQIVNVDPDNLPLDVLVNSCGGIDRC 855
 Db 782 VDASELKPYTRYKRSYSDGFPVKSQDLLEIDLIEHHKVKLVKRVNPDNLVSDTYPDSCSGINRC 841
 Qy 856 SEQQYDANLALNNGCB-NGNMSDSHAPSHIDTGEIDLNENTGIWVVKIPITNGYAT 914
 Db 842 QBOQMVAQLETHHHHPMDCCRAAQTHFHSYIDTGLNSSVDQGIWAIFKVRITDGYAT 901
 Qy 915 LGNLELVEEGPLSGETTLERAAQQOQWQDKMARKGASEKAYYAAQKQIDRLFADYQDQK 974
 Db 902 LGNLELVEGVLGSESLERQDRNTKSAELGRCAETDRVYQDAKQSIHLFVDYQDQ 961
 Qy 975 LNSGVMSDLAAQNLVOSIPYVYNDALPEIPGANYTSFTBLTRLQAMNLYDLRNAIP 1034
 Db 962 LNPEIGHMADIMDAQNLVASISDVYSDAVLQIPGINYEIYETLSNRLQQAQSYLYTSRNVQ 1021
 Qy 1035 NGDFRAGLSDWNTSDVNVQQLSDTSVLVPPNNSQVSOQFTQPNRYVYVLRVARTARKEGV 1094
 Db 1022 NGDFNGLDGNWNTAGASVQDQGNTHFLVLSHWDQAQVSSQPRVQPNCKYVLRVARTARKEGV 1081
 Qy 1095 GDGVYIIRDGANOTETITFNICDDDTGVLSDAQTSYITKTVETPSTBQVWIDMSBETGV 1154
 Db 1082 GDGVYIIRDGANHTETITFNAACDYDINGTYVTDNYLTKKEVIFVTSHTEDMVEVNETBGA 1141
 Qy 1155 FNIESVELVLEE 1167
 Db 1142 FHIDSIEFVETEK 1154

RESULT 4
 S00873
 parasporal crystal protein cryBa1 - Bacillus thuringiensis subsp. thuringiensis
 N;Alternate names: parasporal crystal protein cryA4
 C;Species: Bacillus thuringiensis subsp. thuringiensis
 C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 01-Dec-2000
 C;Accession: S00873
 R;Brizzard, B.L.; Whiteley, H.R.
 Nucleic Acids Res. 16, 2723-2724, 1988
 A;Title: Nucleotide sequence of an additional crystal protein gene cloned from Bacillus
 A;Reference number: S00873; MUID:88203216; PMID:3362680
 A;Accession: S00873
 A;Molecule type: DNA
 A;Residues: 1-1228 <BRI>
 A;Cross-references: EMBL:X06711; NID:940264; PIDN:CAA29898.1; PID:9580949
 C;Genetics:
 A;Gene: cryA4
 A;Start codon: TTG
 C;Superfamily: parasporal crystal protein
 C;Keywords: delta-endotoxin

Query Match 42.1%; Score 2547.5; DB 2; Length 1228;
 Best Local Similarity 43.8%; Pred. No. 1.3e-139;
 Matches 558; Conservative 183; Mismatches 381; Indels 151; Gaps 21;

Qy 1 MSPNNQYEYELDASSSTSVSDNSVRYPLANDQTTTLQNMNYKYDLRMSSEGNPELFGNP 60
 Db 1 MTSNRKNEEIIAVSNHSA-----QMDLLPDARIEDSLCTAEG-----NNI 42
 Qy 61 ETPISSTVGTGIGVQVGLGALVPAGQIASFTSPVIGQVLPSPSTVSVNEMIMKQVED 120

Db 43 DPFFSASTVQTGINIAGRIILGVLPFAGQASFYSLVGLWLPGR-DOMEIIFLHVQ 101
Qy 121 LIDOKITDSVKRTALAGLQGLDGLVYQKSLKWLNRNDTRASVVVTVQVIALELDFV 180
Db 102 LINQITENARTALARLQGLGDSFRAQQSLEDWLENRDDARTSVLYTVQVIALELDFL 161
Qy 181 AKIPSPALSGQEVPLLSVYQAANLHLLLDASIFGAEWFTPGEBISTFYDQVTRTAQ 240
Db 162 NAMPLFAIRNQEVPLLVYQAANLHLLLDASIFGSEFGLTSQEIQRVYQVTRD 221
Qy 241 YSDYCVKMYNTGLDKLKGTAASMLKHOPRREMTLLVLDLVALPNNVDTRTYPIETTAQ 300
Db 222 YSDYCVKMYNTGLSLRGTNAASWRVYQFRRDLGLVLDLVALPNNVDTRTYPIETTAQ 281
Qy 301 LTREYVTDPIVFNRETSGGFCRRSLNSDIISFSEVASVIRSPHLDILSIEFYTRAG 360
Db 282 LTREYVTDPIVFNRETSGGFCRRSLNSDIISFSEVASVIRSPHLDILSIEFYTRAG 337
Qy 361 LPLNTEVLEYVWGHISIKYK-----NTNASSALERNYGTITSNKIKYDLANKDIFQVR 414
Db 338 SRWSNTRHMTYWRGHTIQSRPIGGGLNTSTHGATNTSINPVTLR-----FASRDVYRTE 391
Qy 415 SLGADL--ANYAAOVYGVYPAFSTLLDKNTGSGVGGFTYKSPHTTMOVCTONVNTIDBI 472
Db 392 SYAGVLLWGIYLEPIHGVPTFRFNTFNQNTSDR-GTANYQPYESPGLQKDSST--EL 448
Qy 473 PPE--NEPLSRGYSHRLSHITSYSPSKNASPARYGNLFPVATWTHRSADVNTVYSDKIT 530
Db 449 PPETTERPNYSYSHLSHIGILQSR-----VNVVYVSWTHRSADRTNIGPNRIT 500
Qy 531 QIPVKAHLVSGTIVYKPGFTGQGNILKRTSSGPLATYSVSPLSQRTARIRYAST 590
Db 501 QIPVKAHLVSGTIVYKPGFTGQGNILKRTSSGPLATYSVSPLSQRTARIRYAST 560
Qy 591 TNLRLPVTISGTRIVSINVNTKNGKDDLTFNTPLDLAGITGTAFTPSNYSDSLTVGADSPA 650
Db 561 VDFPFVSRGTVNNFVFLNTSGDBELKYNFVRRAFTPTPTTQIDIRISIQGLS 620
Qy 651 SGGEVYVDKFLIPVATFEAEDLDVAKQAVNGLFTSKD-ALQTSVTDYQVQAANLV 709
Db 621 GNGEYVIDKIEIPVATFEAEDLDVAKQAVNGLFTSKD-ALQTSVTDYQVQAANLV 680
Qy 710 ECLSDLYPNKRMWDVAKRLVQARNLLQDTGPNRIN----- 750
Db 681 ACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPNFTSINKQPDFISTNBQSNFTSIHE 740
Qy 751 -GNGWTSCTGIEVABGDLKORSLRSTAREIDTETPTLYVQOIDESELLKPYRYKL 809
Db 741 QSEHGWGSENITIQEGNDVFAKNYVTLPGT---FNECYPTLYVQKIGESSEKAYRYQL 797
Qy 810 KGFIGSSQDLBKILRHRANOIVKNVP--DNLLPDVL--PVNSCGGIDRCSEQQVVDANL 865
Db 798 RGYIEDSQDLIEYLIRYNAKEHTLDVPGTESLWPLSVESPIGRCEPNRCA--PHFEWNP 855
Qy 866 ALNENGNG--NWSDSHAFSHPIDTGEIDIANENTGIWVVKIPTNGVATLGNLELVEEG 924
Db 856 DLDCSCRDGKCAHSHRFLSDIDVGCCTDLHENLGVVWVFKIKTQEGHARLGNLEFIEEK 915
Qy 925 PLSGETTLERAOOQOOWKMKRKGSEKAYAAKQAIIDLFPADYQDKLNSGVMSMDM 984
Db 916 PLLGALSERVKAERKAKKREKLEKLETRVYVTEAKEAVDALPVDSDYDRLOADNTIGWI 975
Qy 985 LAAQNLVOSIPVYNDALPEIPGMNYTSTFELTNRLQOAMNLYDLRNALPNGDFRNGLSD 1044
Db 976 HAADKLVHRIREAYLSLFPVPGVNAEIPFEELEGHIIITAIISLYDARNVVKNGDFRNGLTC 1035
Qy 1045 WNAISDVNVQQLSDTSVLVIPNWSQVQOFTVQPNRYVLRVPTARKEGVGDVYIRDG 1104
Db 1036 WNVKGVDPVQOSSHRSDLVPEWEAEVSAVRVCPGCGYILRVATYKRGYEGGCVTIHEI 1095
Qy 1105 ANOTETLTF----- 1113
Db 1096 ENNTDELKFKRHEEVPTDTGTCTNDYTAHQGTAGCADACNSRNAGYEDAYEVDTTASV 1155

Qy 1114 -----NICDDTGVLSAD--QTSYITKTVEFTSTBQVWIDMETBQV 1154
Db 1156 NYKPTYEBETVDRDNNHCYDRGVVNPVPPAGVYVTKELFYFPETDTVWIEIGTEGK 1215
Qy 1155 FNIESVELVLEEE 1167
Db 1216 FIVDSVELLMEE 1228
RESULT 5
A48944
paraaporal crystal protein cry7Aal - Bacillus thuringiensis
N:Alternate names: paraaporal crystal protein cryIIIC
C:Species: Bacillus thuringiensis
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C:Accession: A48944
R:Lambert, B.; Hofte, H.; Anny, K.; Jansens, S.; Soetaert, P.; Peferoen, M.
Appl. Environ. Microbiol. 58, 2536-2542, 1992
A>Title: Novel Bacillus thuringiensis insecticidal crystal protein with a silent activi
A:Reference number: A48944; MUID:92384571; PMID:1514800
A:Contents: B7S1373
A:Accession: A48944
A>Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-1138 <LAM>
A:Cross-references: GB:M64478; NID:g142760; PIDN:AAA22351.1; PID:g142761
A>Note: sequence extracted from NCBI backbone (NCBIN:112092, NCBIPI:112093)
C:Superfamily: paraaporal crystal protein
C:Keywords: delta-endotoxin

Query Match 40.7%; Score 2458.5; DB 2; Length 1138;
Best Local Similarity 43.5%; Pred. No. 1.6e-134;
Matches 514; Conservative 208; Mismatches 400; Indels 59; Gaps 17;

Qy 1 MSPNNQNEYILDASSSTVSNSVRYPLANDQTTTLQNMNYKDYLRMSGEBNPELFGNP 60
Db 1 ANLNLDGYS-----DSNRTLANSNLYPTQKALSPLKMNQYDFLSITRQEPALASG 55
Qy 61 ETTFSSSTVQTGIGIVGQVLAGVPFAGQIASFYFVIGQLWPSSTVSVWEMKQVED 120
Db 56 NT-----AINTVSVTVGATLSALGVGASFTNPKYLIAGLLWPENG-KIMDFNTEVEA 109
Qy 121 LIDOKITDSVKRTALAGLQGLDGLVYQKSLKWLNRNDTRASVVVTVQVIALELDFV 180
Db 110 LIDOKIEEYVNRKAIELDGLGSLADRYQKALADWLKQDDPEALISVATPERIIDSLPE 169
Qy 181 AKIPSPALSGQEVPLLSVYQAANLHLLLDASIFGAEWFTPGEBISTFYDQVTRTAQ 240
Db 170 FSMSPKVTGYEIPLLTVYQAANLHLLLDASITLYGDKNGFTQNNIBENTYRQKRIS 229
Qy 241 YSDYCVKMYNTGLDKLKGTAASMLKHOPRREMTLLVLDLVALPNNVDTRTYPIETTAQ 300
Db 230 YSDCTKWNYSGLSRNLSTYEQVNTYNRFRREMLMALDLVAVPPFDPFRYSMETSTQ 289
Qy 301 LTREYVTDPIVFNRETSGGFCRRSLNSDI--SSEVESAVIRSPHLDILSIEFYTR 358
Db 290 LTREYVTDPIVFNRETSGGFCRRSLNSDI--SSEVESAVIRSPHLDILSIEFYTR 338
Qy 359 AGLPLNTE-YLEYVWGHISIKYKNTNASSALERNYGTITSNKIK--YVDLANKDIFQVR 415
Db 339 YKASHEIOPDLFWSAHKVSFKSEQSNLYTGYGTSGYISSGAVSPHGNIDIVRTLA 398
Qy 416 LGADLANYYAQVGVYPAFTLLDKNTGSGVGGFTYKSPHTTMOVCTONVNTIDPE 475
Db 399 APSVWVYPTQNYGVQVEFY-----GVKGVHVRGDKNYDL---YDSDIDQLPPD 446
Qy 476 NEPLSRGYSHRLSHITSYSPSKNASPARYGN--LPVATWTHRSADVNTVYSDKITQIP 533
Db 447 GEPHIEKTHRLCHATAI-----FKSTPDYDNATIPFSWTHRSABYRNIRPNKITKIP 501
Qy 534 VVKAHTLVSGTIVYKPGFTGQGNILKRTSSGPLATYSVSPLSQRTARIRYASTTNL 593

Db 502 AVKMYKLDPPSTVVKPGFTGCDLVKRGSTGYIGDIKATVNSPLSQYRVVRVATNVSG 561
Qy 594 RLFTTISG-----TRIYSINNVKTMNGKDDLTFTNFDLATIGTAFTSNTSDSLTVGADS 648
Db 562 QFNYYINDKITLQTKFC--NTVETIGEGKDLTSGFGYIYSTTIOFPDEPKITLHSD 619
Qy 649 FASGGEVYVDFELIPNATPEABEDLDVAKKAVNGLPTSKDQALQTSVTDYQVQNAAL 708
Db 620 LSNSSFYVDSIFPIPDVNVYAEKEKLEKAKAVNTLFTTEGNALQNDVDYKQVDSIL 679
Qy 709 VECLSDLEYPNKREMLWDVAKREKRLVQARNLLQDTGFNRING--ENGWGTGSGTIEVAEG 766
Db 680 VDCISGDLYPNKKELQNLVYAKRLSYRNMLLDPTFDSINSEENGWYSGNGIVLNG 739
Qy 767 DVLFKORSKLTSAEIDTETPTLYLQQIDESLLKPTTRYKLGKPGTSSQDLLEIKLRH 826
Db 740 DFVFKGNLYLIFSGTN--DTQ-YPTLYQKIDESKLEKKEYTRYKLGKFTESSQDLLEAVYR 796
Qy 827 RANQIVKQVNDLLPDVLPVNSCGIDRCSEQQVVDANLALNNGENGNSSDSHAFSPH 886
Db 797 DAKHRTLDVNDLLPDILPENTCEPNKCAQQYLDENPSPSCSMQDGLSDSHSPSLN 856
Qy 887 IDTGEIDLNENTGIWVFKIPITNGYATLGNLELVEBGLSGTETLERAQQEQEQODKMA 946
Db 857 IDTGSINHNENLGIWLFKISTLEGYAKFKNLEVEBGPVIGALARVKRQETKRNKLA 916
Qy 947 RKGAASEKAYAAKQALIDRLPADYQDQKLSGVEMSDMLAAQNLVQSPVYNDALPEIP 1006
Db 917 QLATTETQALYTRAKQALDNLFANAQDHLKRDVTFABIAARKIVQSIREAYMSLWVVP 976
Qy 1007 GMYVTSFTELNLQOAWNLVDLNAIPNGDFRGLSDNATSDVNVQQLSDTSVLVPIPN 1066
Db 977 GVNHPIFTELGRVQRAFQYDVNVNVRNGRFLNGLSDMVITSDVRQBEKNGNVLVNN 1036
Qy 1067 WNSQVSQOFTVQPNRYRVLRTARKEGVDGYIIRDEGANQETLTLPNICDDDTGVLSD 1126
Db 1037 WDAQVLQNLVLYQDRGYLHVLTARKIGIGBYITIDEEGHTDQLRFTAC-BEIDASNAF 1095
Qy 1127 QTSYITKTVPTSTEQVMDMSETEGVNFVNSVELVLEE 1167
Db 1096 ISGYITKLEFFPDTEKVHIEIGTEGIFLVESIELFLMBE 1136

RESULT 6
S19306
paraaporal crystal protein cry9Aa1 - Bacillus thuringiensis
N;Alternate names: delta-endotoxin; insecticidal crystal protein; paraaporal crystal pro
C;Species: Bacillus thuringiensis
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 01-Dec-2000
C;Accession: S19306; S23588; A44847; S14602; S14837
R;Smulevitch, S.V.; Osterman, A.L.; Shevelev, A.B.; Kaluger, S.V.; Karasin, A.I.; Kadyrc
FBBS Lett. 293, 25-28, 1991
A;Title: Nucleotide sequence of a novel delta-endotoxin gene cryIG of Bacillus thuringie
A;Reference number: S19306; MUID:92070568; PMID:1660003
A;Accession: S19306
A;Molecule type: DNA
A;Residues: 1-1156 <SMU>
A;Cross-references: EMBL:X58120; NID:9870929; PIDN:CAA41122.1; PID:g40271
A;Experimental source: subsp. gallerae
A;Accession: S23588
A;Molecule type: protein
A;Residues: 24-34 <SMU>
A;Experimental source: subsp. gallerae
R;Gleave, A.P.; Hedges, R.J.; Broadwell, A.H.
J. Gen. Microbiol. 138, 55-62, 1992
A;Title: Identification of an insecticidal crystal protein from Bacillus thuringiensis
A;Reference number: A44847; MUID:92211329; PMID:1556556
A;Accession: A44847
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1151 <GLX>
A;Cross-references: EMBL:X58534; NID:g48879; PIDN:CAA41425.1; PID:g48880
A;Experimental source: isolate DSR517

A;Note: sequence extracted from NCBI backbone (NCBIN:92865, NCBIP:92867)
C;Genetics:
A;Gene: cryIG
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 36.3%; Score 2191; DB 2; Length 1156;
Best Local Similarity 41.5%; Pred. No. 5.7e-119; Indels 98; Gaps 24;
Matches 501; Conservative 195; Mismatches 414;

Qy 5 NQNEYELDASSSTVSVDNSVRYPDLANDQTTTLQNMNYKQYLRMSEGENPBLFCGNPETPI 64
Db 2 NQNKGGIIGASNCOCASDDVAKYPLANNYPSSALNIN-----SC 40

Qy 65 SSSTVQTGIGIVGQ-----VLGALGVP-FAGQIASFYSPVQQLPSPSTVSWE 112
Db 41 QNSILANWINIIGDAKAEAVSIGTIVSLITAPSLTGLISIVYDLIGKVLGSGSQSISD 100

Qy 113 MIMKQVEDLDQKLTDSVRKTLAAGLQGLDGVYQKSLKNWLENRNDTRARSVVVYQY 172
Db 101 LSIQDLUSIIDLRVSQSVLNDGIADFNQSVLLYRNYLEALDSWKNKNPNSASABELRTRPR 160

Qy 173 IA-LELDFVAKIPSPFAISGQ-----EVPLLSVYAQAANLHLLLRDASIFGAEMGP--- 222
Db 161 IADSEFDRLTRGSLTNGSSLARQNAQIILLPSASAAPFHLILLRLDRATRYGTWGLYNA 220

Qy 223 TPGEISTPYDROVTRTAQYSDYCVKWTNTGLDKL--KGTNAASWLYKHQPREMTLLVLD 280
Db 221 TP--PINYQSKVELIBLYTDYCVHMYNRGPNELRQRTSATAWLEPHRYREMTLMVLD 278

Qy 281 LVALPNYDTRTPIETTAQLTREVVYDPIVFNRETSGPCRRWSLNSD-----ISP 332
Db 279 IVASFSSLDITNTPIETDFQLSRVYITDPI-----GPHRSLSRGESWFSFVNRANP 330

Qy 333 SYVESAVIRSEPHLFDILSIEFYTTTRAGLPLN-NTEYLEYVWGHISIKYKTNASSALERN 391
Db 331 SDLENA-IPNPRSPFWLNNMIISGSLTLPVSPSTDRARVWYSGRDRISPANSQPIBELI 389

Qy 392 YGITSNKIKYDOLANKDIFQVRSGLADLANYYAQVGVYPYASFTLDDKQYGS-----G 445
Db 390 SQQHTT---ATQITLGERIFRVDQSQCANLD---TTYGVNRAVP-YHDASGSGRSRYVEG 442

Qy 446 SVGGFTYKPHHTMQVCTQNYNTIDEIPPEE--PLSRGYSHRLSHITSYFSKNASPA 503
Db 443 YIRTTGIDNPR-----VQNTY--LPGENSDIPTPETHILSTINLTGGLRQVASN 494

Qy 504 RYGNLPVFAWTHRSADVNTVYSDKITQIPVKAHTLVSGTTVTKGPGFTGKNILKRTSS 563
Db 495 RRSSLVWYGTWTHKSLARNNTINPDRIQIPLTKVDTRGTGVSYYVNDPGFIGALLQRTDH 554

Qy 564 GPLATYSVSVKSPLSQRYRARIYASTNRLPFTIS-GTRIYSINNVKTNKGGDLTPN 622
Db 555 GSLGVLRVQPLHLRQOYRIRVYASTNRLSVNGSFGT--ISQNLPTNRLGEDLAYG 612

Qy 623 TFDLATIGTAFPSNYSDSLTVGADSPASGGEVYVDKPELIPVNAATPEABEDLDVAKKAV 682
Db 613 SPAIREPNTSIRPTASPDQRLFTIETPSFIRQEVTVDRIBFIPVNPTRAKEDLEAKKAV 672

Qy 683 NGLFTSKDQALQTSVTDYQVQNAALVBCLSDELYPNEKMLWDVAKREKRLVQARNLLQ 742
Db 673 ASLPTTRDGLQVNVKDYQVDQAANLVCSLSDEQYGVYDKMLLEAVRAAKSLSRNLLQ 732

Qy 743 DTGFNRING--ENGWGTGSGTIEVAEGDLFKDLSRLTSAREIDTETPTLYLQQIDESL 800
Db 733 DPDEFNTINSTEENGWAKASGVITSEGGPFYKRAIQLASAR-----ENYPTIYQKVDASE 788

Qy 801 LKPYTRYKLGKFTGSSQDLLEIKLRHANQIVKQVNDLLPDVLPVNSCGGIDRCSEQY 860
Db 789 LKPYTRYRDLGFKVSKQDLLEIKLRHANQIVKQVNDLLPDVLPVNSCGGIDRCSEQY 848

Qy 861 VDNALALENNGE-NGNMSDSDSHAFSPHIDTGEIDLNENTGIWVFKIPITNGYATLGNLE 919
Db 849 VNAQLETEHEHPMDCCCEAAQTHFSSVIDTGDNLSSVDQGIWAIFKVRTTDDGYATLGNLE 908

Query Match 34.6%; Score 2093.5; DB 2; Length 1166;
Best Local Similarity 39.7%; Pred. No. 2.7e-113;
Matches 482; Conservative 172; Mismatches 392; Indels 167; Gaps 24;

QY 59 NPETFISSVTQIGIVGOVLGAL-----GYPEAGQIASFYFVGOGLWPSSTVSVM 111
DB 17 NPSEIFNAR-NSNFGVLVSQSSGLTRPFLLEAAVPEAGFALGLFDII-----WGLGVDQW 71

QY 112 EMIMKQVEDLDQKITDSVRTALAGLDGDDVYKSLKNLWNRNDRASVVVTQ 171
DB 72 SLFLRQIEQLRQITELERNRATAILTGLSSSNLYVEALREWENDPNNPASQVRTR 131

QY 172 YIALELDPVAKIPSAISQGVPLLSVYQAANLHLLLRDASIFGAEWGETPGEISTFY 231
DB 132 FRUTDDAIVTGLPFLAIRNLEWNLISVYTOAANLHLLLRDASVYFGRWGLTQANIEDLY 191

QY 232 DRQVTRTAQYSDYCVKMYNTGLDKLKGNTAASWLYKHQFRREMTLLVLDLVALFPNYDTR 291
DB 192 TRLTNSIOEYSDHARWYNQGLNEIGISR-----RYLDFQDRLTISVLDLVALFPNYDIR 247

QY 292 TYPIETPAQLTREYVTDPIVFNRETSGGFCRRWSLNSDISFSVESAVIRSPHFLDILSE 351
DB 248 TYPIETQSLTREYTSFVV-----AGNINFGLSIANVLRAPHLMDPIDR 292

QY 352 IEPTTTRAGLPLNTEYLEYVWGHISIKYKNTNASSALERN---YGT----- 394
DB 293 IVIYT-----NSVSTPYWAGHEVISRRTGQGGNEIRFPLVYGAANAABPPVITIRPTG 345

QY 395 ITSNIKIKYDLANKDIFQVRSGLADLANVYQVYGVYASFTLMDKNTGSGVGFT-YS 453
DB 346 FTDEQRQWY-RASRVVVSFRSSGQD-----FSLVD-----AVGFLTIFS 383

QY 454 KPHTTMQVCTONYNTIDEIPPENEPLSRGYSRLSHLTSYSPSKNASSP--ARYGNLPVP 511
DB 384 AVSIYRNGFGFNTDIDIEPIEGTDPPTGYSHRLCHVGFL-----ASSPPTISQVAPAP 438

QY 512 ANTHRSADVNTVYSDKITQIPVVKAAHTLVSGTIVKGPQFTQGNILKRTSSGPLATSV 571
DB 439 SWTHRSATLNTIAPDVIITQIPKAFNLHSGATIVKGPQFTGCDILRRTNVGSGDMRV 498

QY 572 SVKSPLSORVRYRYASTTNLRFLVTISGTRIVSYINVNTKMGDDLTNTFDLATIGT 631
DB 499 NITAPLSORVRYRYASTTDLQPTYNINGTTINGNFSTMGDDLOQYGRFVAGFTT 558

QY 632 APTFSNYSDSLTVGADSFASGGEYVDKFLIPVNATFEABEDLDVAKAVNGFLTSGD 691
DB 559 PPTFSDANSTFTIGAFGSPNNEVYIDRIEVPVPAEVTFEARYDLEKAKAVNALFTSSNQ 618

QY 692 -ALQTSVTDYOVQOANLVECLSDLYPNKRMMDAVKAKRLVQARNLLQDTGNRIN 750
DB 619 IGLKTDVTDYHDKVSNLVECLSDLYPNKRMMDAVKAKRLVQARNLLQDTGNRIN 678

QY 751 --GENGWTSGLIYVAGSDVLEKDRSLRLTSAREIDETPTLYLQOIDSLLKPYTRYK 808
DB 679 RQDRGRGSDTITQGGDVFKEVNYVLPCTFD---GCYPTLYLQIDSKLKVITYYQ 735

QY 809 LKGPFGSSQDLKILIRHRANQIVKVPDN-----LLPDVLPVNSCGGIDRCSEQQYVDAN 864
DB 736 LRGYEDSODLEIYLIRYNAKHETVNVPGTGLWPLSAQSPIGKCGEPNCA--PHLEWN 793

QY 865 LALENGENG-NMSSDSHAFSFHIDTGEIDLNENGTGWVFKIPTTNGYATLGNLELVRE 923
DB 794 PDLDCSRNCEKCAHSHHESLDIDVGCTDLNEDLGVWVIFKIQDGHARLGNLEPLEE 853

QY 924 GPLSGETLERAQQEQOQWQMARKRGASEKAYAAQADRLPADYQDQKLSGVEMSD 983
DB 854 KPLLGALARYKRAEKKWRDREKLELETINIVYKEASEVDALFVNSQYDOLQADTNIAM 913

QY 984 MLAANLVQSIPIVYNDALPEIGPMYNTSFTELNRQQAWNLVLDENAIPNGDFRNLIS 1043
DB 914 IHAADKKVHSIREAYLPVLPVGNAAI FEELEGRIFTAFSLYDARNVINGDFRNLIS 973

1044 DMNATSDVNV-QQLSDTSVLVIPNNNSQVSOFTVQPNRYVLRVYTABKEGVGDGVIR 1102
DB 974 CMNVKGVHDVEEQNNHRSVLVPEWBAEVSQVPCGCGHILRTAYKEGVGSCVTH 1033

QY 1103 DGAQOTETLTPN-----ICDDDTG-----VLSA 1125
DB 1034 EIEENTDELAFSNCVBEVYPNMTVTCDNYTANQEEYKGAVTSHNRGYDEAVGNPSPVA 1093

QY 1126 DOT-----SYTKTVETPSTQWIDMSBTEGV 1154
DB 1094 DYTTPVEEKAYTDGRRENPCBSNRGSDYDTPPLPAGYVYTKLEYEYPTDKVWIEIGETST 1153

QY 1155 FNIESVELVLEE 1167
DB 1154 FIVESVELLME 1166

RESULT 9
A26513
parasporal crystal protein - Bacillus thuringiensis (strain aizawai)
C;Species: Bacillus thuringiensis
C;Date: 11-Mar-1988 #sequence_revision 11-Mar-1988 #text_change 01-Dec-2000
C;Accession: A26513
R;Oeda, K.; Oshie, K.; Shimizu, M.; Nakamura, K.; Yamamoto, H.; Nakayama, I.; Ohkawa, H.
Gene 53, 113-119, 1987
A;Title: Nucleotide sequence of the insecticidal protein gene of Bacillus thuringiensis
A;Reference number: A26513; MUID:87248103; PMID:3297927
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1155 <OED>
A;Cross-references: GB:M16463; NID:g143098; PIDN:AAA22551.1; PID:g143099
C;Superfamily: parasporal crystal protein
C;Keywords: Delta-endotoxin

Query Match 34.2%; Score 2064.5; DB 2; Length 1155;
Best Local Similarity 39.0%; Pred. No. 1.3e-111;
Matches 476; Conservative 184; Mismatches 387; Indels 173; Gaps 27;

QY 53 NPGL-----FENPET-FISSVTQGTG---IGI---VGQVLGALGVPPAGQIASFYSP 97
DB 4 NPINECIPYCNLSNPBEVLGGERIETGVTPIDLSLSTQFLSEFPVAGFVLGLVDI 63

QY 98 IVGQLWPSSTVSVMEMKQVEDLDQKITDSVRTALAGLDGDDVYKSLKNLW 157
DB 64 IWGIFGFSQ---WDAFLVQIEQLINQRIEFPARNQAIISRLGSLNLYQIYAESFRWEA 119

QY 158 NRNDTRARSVVVTOYIALELDFVAKIPSAISQGVPLLSVYQAANLHLLLRDASIFG 217
DB 120 DPTNPALREMRQFNDMSALITAIPLFAVQVQVPLLSVYQAANLHLSVLRDVSYPG 179

QY 218 AEMGFTPEISTFYDQVTRTAQYSDYCVKMYNTGLDKLKGNTAASWLYKHQFRREMTLL 277
DB 180 QRWGFDAAATINSRYNDLTILGNVTDHARWYNGLERWVGDPDSRDWIRYVQFRRELTLT 239

QY 278 VLDLVALFPNYDTRTYETIETTAQLTREYVTDPIVFNRETSGGFCRRWSLNSDISFSVES 337
DB 240 VLDIVSLFPNYDSTRYETIETVIRTYSQLTREIYTNPLV-----ENPDGFRALAQ 285

QY 338 AV---IRSPHLFDLSIEFVYTRAGLPLNTEYLEYVWGHISIK-----YKNTNASSALE 389
DB 286 GIGSIRSPHMLDILNSITVT-----DAHGEY--YWSGHQIMASPVGSPGPTFFL- 337

QY 390 RNYGTITSNIKIYVDLANHDIPOVRSIGADLANVYQVY--GVYASFTLLDKNT-----G 443
DB 338 --YGTMGNAAPQQRIVQAQLCGQVYVRLTSLT---YRRPFNIGINNQQLSVLDGTGFAYGT 392

QY 444 SGSVGGTYSKPHHTMQVCTQNTNIDEIPEPE--PLSRCYSRLSHLTSY--SPSKNA 499
DB 393 SSNLPAVYKSGTV-----DSLDEIPQNNVPPROGFSRLSHSVNFRSGFSSNS 444

QY 500 SSPARYGNLVPFAWTHRSADVNTVYSDKITQIPVVKAAHTLVSGTIVKGPQFTGGLK 559

Db 445 VSIIR--APFWSIHRSAEWNIIIPSSQITQIPLTKSTNLSGTSVVGKPGFTGDIIR 501
Qy 560 RTSGPLAYTSVSKSPQSRYRARIYASTNLRFLFTISGTIRYSINNVKTNKGGDL 619
Db 502 RTSPQISTLRVNTAPLSQRVIRYASTNLRFLFTISGTIRYSINNVKTNKGGDL 561
Qy 620 TFMFDLATICTAPTESNYSLSLVGADSPASGGEVYVDKELIPVNAATFEARBDLVAK 679
Db 562 QSGSFRVTGFTPPNFSNGSVFTLSAHVFNSENEVIDRIEFVPAEFTFEAYDLERAQ 621
Qy 680 KAVNGLFTSKD-ALQTSVTDYQVQANLVKLSDELPEYNEKRLMDAVKAEARLVQAR 738
Db 622 KAVNELFTSSNQIGLKTVDYDHIQVSNLVECLDFECLDEKSEKVKHAKLSDER 681
Qy 739 NLLQDTGPNRNG--ENGWTSCTGIEVABGVLPKRSRLRSTARSABIDTPTVLYQOI 796
Db 682 NLLQDPNFRGINQLDRGWRGSDTITIQGGDDVFKENYVTLTGTFD--ECYPTVLYQKI 738
Qy 797 DESLLKPVTRKLGFTSSODLEIKLRHQAQIVQVDPDNLAPDVLVNVSCGIDRCS 856
Db 739 DESLLKPVTRKLGFTSSODLEIKLRHQAQIVQVDPDNLAPDVLVNVSCGIDRCS 796
Qy 857 EQQVVDANLALNENGENGSSSDSHAFSHDITGIDLNENTGIWVVKIPTTNGYATLG 916
Db 797 HH-----SHFSLDIDVGCTDLNEDLGWVVIKIKTQDGHARLG 835
Qy 917 NLELVREGLSGETLERAQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 976
Db 836 NLELVREGLSGETLERAQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 895
Qy 977 SGVMSDMLAQNVLQVSTPYVNDALPEI PCGNYTSFTLNRLOQANVLDLNAIPNG 1036
Db 896 ADTWIAMIHADKRVHSIREAYLPELSVPGVNAIPELSEGRIFTAFSLYDARNVKNK 955
Qy 1037 DFRNGLSDNATSDVNV--QLSDTSVLVPIPNNSQVSOQFTVQPNRYVLRVTRARKEGV 1095
Db 956 DFNGLSCMVVKHVDVEEQNQSRVSVLPEWEARVSVQVPCRGVILAVTAYKEGV 1015
Qy 1096 DGYVIRGQANQVETLTN-----ICDDTGV----- 1122
Db 1016 EGCVTIHEIENNTDELKFSNFVEEVEYNNVTVCNDYTATQSEYEGTYSNRNGDYGE 1075
Qy 1123 ----LSADQTS-----YIKTVEFTSPSTQVWID 1147
Db 1076 SNSVPADYASAYEKATDGRDNDPCSNRGYGDYTPLPAGYVTKELIYFETDKWIE 1135
Qy 1148 MSETGVENIESVHVLBEE 1167
Db 1136 IGETEGTIVDSVLLAMEE 1155
RESULT 10
JD0002
Parasporal crystal protein crylab3 - Bacillus thuringiensis
N:Alternate names: delta-endotoxin-2; entomocidal protoxin; entomopathogenic crystal pro
C:Species: Bacillus thuringiensis
C>Date: 28-Dec-1987 #sequence revision 28-Dec-1987 #text change 01-Dec-2000
C:Accession: A30025; A91560; A90955; S14555; A26461; A24172; A29043; JD0002
R:Kondo, S.; Tamura, N.; Kunitate, A.; Hattori, M.; Akashi, A.; Ohmori, I.
Agric. Biol. Chem. 51, 455-463, 1987
A:Title: Cloning and nucleotide sequencing of two insecticidal delta-endotoxin genes fro
A:Reference number: A90025
A:Accession: A90025
A:Molecule type: mRNA
A:Residues: 1-1155 <KON>
A:Experimental source: subsp. kurstaki
R:Geiser, M.; Schweitzer, S.; Grimm, C.
Gene 48, 109-118, 1986
A:Title: The hypervariable region in the genes coding for entomopathogenic crystal prote
A:Reference number: A91560; MUID:87163505; PMID:3557124
A:Accession: A91560
A:Molecule type: DNA
A:Residues: 1-1155 <GEI>

A:Cross-references: GB:ML5271; NID:G143123; PIDN:AAA22561.1; PID:G143124
A:Experimental source: subsp. kurstaki
R:Wabiko, H.; Raymond, K.C.; Bulla Jr., L.A.
DNA 5, 305-314, 1986
A:Title: Bacillus thuringiensis entomocidal protoxin gene sequence and gene product ana
A:Reference number: A90955; MUID:86300092; PMID:3743328
A:Accession: A90955
A:Molecule type: DNA
A:Residues: 1-1155 <WAB>
A:Cross-references: GB:ML3898; NID:G142719; PIDN:AAA22330.1; PID:G142720
A:Experimental source: subsp. berliner
R:Chak, K.F.; Jen, J.C.
submitted to the EMBL Data Library, October 1990
A:Description: Complete nucleotide sequence and expression in Escherichia coli of a cry
A:Reference number: S14555
A:Accession: S14555
A:Molecule type: DNA
A:Residues: 1-1155 <CHA>
A:Cross-references: EMBL:X54939; NID:G40272; PIDN:CAA38701.1; PID:G40273
R:Hofte, H.; de Greve, H.; Seurinck, J.; Janssens, S.; Mahillon, J.; Ampe, C.; Vandekerck
Eur. J. Biochem. 161, 273-280, 1986
A:Title: Structural and functional analysis of a cloned delta endotoxin of Bacillus thu
A:Reference number: A26461; MUID:87054026; PMID:3023091
A:Accession: A26461
A:Molecule type: DNA
A:Residues: 1-730, L', 732-784, R', 786-1155 <HOP>
A:Cross-references: GB:X04698; NID:G40254; PIDN:CAA38405.1; PID:G40255
A:Experimental source: strain berliner 1715
C:Comment: This toxin is lethal to the larvae of lepidopterans and dipterans.
C:Genetics:
A:Gene: cry-1-2; bt2
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin
P:82-586/Product: toxic peptide #status predicted <TXP>
P:82-300/Region: toxic #status predicted
P:300-586/Region: insecticidal #status predicted
Query Match 34.1%; Score 2061.5; DB 2; Length 1155;
Best Local Similarity 39.3%; Pred. No. 1.9e-111;
Matches 478; Conservative 185; Mismatches 387; Indels 167; Gaps 29;
Qy 53 NPGL-----FGNPET-FISSSTVQVG---IGI---VGQVIGALGVPPAQIASFYSF 97
Db 4 NPNECIPYCNLSNPEVEVIGGERIEGYTIDISLSTQELLSEFPVAGFVLGLVDI 63
Qy 98 IVGQLWPSSTVSVMEMKQVEDLIDQKITDSVRKTALAGLQGLGDLDVYQKSLKMWLE 157
Db 64 IWGIFGPGSQ---NDAPLVQIEQLINQRIEPPARNOAISRLGLSNLYQIYAESFWEA 119
Qy 158 NRNDTRARSVVVTVYIALELDFVAKIPSPALSGQEVPLISVYQAANLHLLLRDASIFG 217
Db 120 DPTNPALREEMRIQFNDMNSALTTAIPLVAVQVQVPLLSVYQAANLHLLSVLDVSFVG 179
Qy 218 AEWGFTPCIEISTFYDQVTRTAQVSDYCVKWTGCLDKLKGTAASLWKYHOFREMTLL 277
Db 190 QRMGDDAATINSRYNDLRLTLLGNVTHAVRWNTGLERWGDSDRWYQVPRRLTLLT 239
Qy 278 VLDLVALPNTYDTRTYPIETTAQLTRVYTPDIVFNRETSGGFCRRWSLNSDISFSEVS 337
Db 240 VLDIVSLFPNYSRTYPIRTVSQLTRVYTPNVLNFD--GSF--RGSAGQ-----IEG 289
Qy 338 AVIRSPHLFDLSEIEFTYTRAGLPLNNTVYLVYVGHSHK-----YKNTNASSALERNY 392
Db 290 S-IRSPHLMQILNSITIT-----DAHGEY---YWSGHQIMASPVGFSQPEFTFPL---Y 338
Qy 393 GTITSNKIKYYDLANKDIFQVRLGADLANYYAQVY--GVYPASFLLDKNT-----GSGS 446
Db 339 GTWGNAAQQRIVAGLQGVYVTLSSLT---YRFPNIGINNQQSLVLDGTETFAVCTSSN 395
Qy 447 VGGFTYKPHHTMQVCTQNTYNTIDIPENE--PLSRGYSHRLSHITSY--SFSKNASFP 502
Db 396 LPSAVYKSGTV-----DSLBIQNNNVPPRQGFPSHRLSHVMSFRSGFSNVS81 447

Db 791 PI-----GKCAHSHHFLSDIDVCTDLNEDLGWVYFKITQDGHARLGNLE 838
Qy LVEEGPLSGETLERAQOQOQOMARKRGASEKAYVAAQAOIIDLPLADYQDOKLNSGV 979
Db 839 FLEKPLVGEALARVRAKAKWRKREKLEMETNIVYKEAKESVDALPVNSQYDLQADT 898
Qy 980 EMSDLAAQNLVQIPYVYNDALPEIPGMNNTSFTLNRLOQANWLYDLNAPNGDPR 1039
Db 899 NIAMIHAADKRVHSIREAYLPELSVPGVNAAIPEELGRIPTAFSLYDARNVKNKGDPN 958
Qy 1040 KGLSDNRATSDVNV-QQLSDTSVLVIPNNSQVSOQFTVQPNRYVLTATKRGVGGY 1098
Db 959 NGLSCWNVKGVHDVEEQNHNSVLVPEWEAEVSOEVKPCRGYILRVATYKGYGEGC 1018
Qy 1099 VIIRGANOTETLTFN-----ICDDTGV----- 1122
Db 1019 VTIIHEINNTDELKESNCVEEVPNNVTTCNDYATQOEYEGYTSNRGVDGAYSNS 1078
Qy 1123 -LSADQTS-----YITKTVETPSTEQWIDMSR 1150
Db 1079 SVPADYASAVEKAYTDCRRDNPCESNRGVDYTPPLPAGYVTKLEYETPDKVWIEIGE 1138
Qy 1151 TEGVFNIESVELVLEE 1167
Db 1139 TEGTFIVDSVELLME 1155

RESULT 12
JT0241
Parasporal crystal protein - *Bacillus thuringiensis* (strain aizawai IPL7)
N:Alternate names: 135K insecticidal protein
C:Species: *Bacillus thuringiensis*
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 01-Dec-2000
R:Shimizu, M.; Oshie, K.; Nakamura, K.; Takada, Y.; Oeda, K.; Ohkawa, H.
Agric. Biol. Chem. 52, 1565-1573, 1988
A:Title: Cloning and expression in *Escherichia coli* of the 135-kDa insecticidal protein
A:Reference number: JT0241
A:Accession: JT0241
A:Molecule type: DNA
A:Residues: 1-1176 <SH>
C:Comment: The 135K protein has insecticidal activity against *Plutella xylostella* larvae
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.9%; Score 2048; DB 2; Length 1176;
Best Local Similarity 39.3%; Pred. No. 1.2e-110;
Matches 480; Conservative 174; Mismatches 414; Indels 152; Gaps 26;

Qy 53 NPET-----FGNPET-FISSSTVQTG---IGI---VGVLGALGVPPFAGQIASFYSP 97
Db 4 NNINECIPNCLSNPEVVLGRIETGYTPIDISLSITQLLSEFPVPGAGFVLGLVDI 63
Qy 98 IVGQLMPSTSVSWEMIMKQVEDLIDQITSVRTALAGLQGLDGLDVTOKSLNMLE 157
Db 64 IWGIFGPSQ---WDAFLVQELINQRIEAPRNQAIISRLGSLNLYQIYAESPRWEA 119
Qy 158 NRNDTRASVVVTVYALELDFVAKI BSPAISQGVPLLSVYAOANLHLLLRDASIFG 217
Db 120 DPTNPALREEMRIQPDNMSALTALPLFAVQNYQVPLSVYVQANLHLSVLRDVSVFG 179
Qy 218 AEWGTPCEISTFYDQVTRTAQYSDYCVKWTNTGLDKLKGNTNAAASMLKTHQFREMTELL 277
Db 180 QRWGPDATINSRYNDLITELIGNYTDYAVRWNTGLERWVGPDSDRDVRYNQFRELIT 239
Qy 278 VLIDLVALFPNVDRTYPIETTAQLTRVYTDIVFNRTS-GGFCRRSLNSDISFSRVE 336
Db 240 VLIDVALFSNDSRRYPITVSQLTREIYTNVLENFPGSFRGMAQRIEQN----- 290
Qy 337 SAVIRSPHLPDILSIEFTT-TRAGLPLNNTLEYLWVGHSTK-----YKNTNASSALER 390

Db 291 ---IROPHMLDILINSIITYTDVHRG-----FYNWSGHQITASPVGSGPEFAPPLPG 339
Qy 391 NYGTITSNKIKYDYLANRDIQVRS-----LGADLANYAYQYGVYVYASFTLLDKRT 442
Db 340 NAGNAAPVL--VSLTGLGIFRTLSSPLRYRIILGSGPNQBELFVLGDTFESFASLTNL 397
Qy 443 GSGSVGGFTYSKPHHTMQVCTQNTNTDEIPENE--PLSRGYSRHSLSHTSYSSKNAS 500
Db 398 PS-----TIYRQGTV-----DSLVIIPQNSVPPFRAGFSRHSLSHTVMSQAAGAV 444
Qy 501 SPARYGNLPVAMTHRSADVTNVYSDKIQTQIPVYKAHTLVSGTTVIKGGPGTGNILKR 560
Db 445 YTLR---APTFSWHRSAEFNNIIPSSQITQIPLTKSTNLGSGTSVWAGPGTGGDILRR 501
Qy 561 TSSGPLATYSVSKSPLSQRYRARIYASTTNLRLFTVITSGTRIYSINNVKTNMGDDLT 620
Db 502 TSPQIISLIRVNITAPLSQRYRARIYASTTNLPHSTSIDQRPIINQGNFSAFMSSGSNIQ 561
Qy 621 FNTFDLATIGTAFPSNYSDSLTVGADSPASGGEVYVDKFLIPVNATFEAREEDLDVAKK 680
Db 562 SGRFRTVGTFTPPFNSGSSVFTLSAHVFNESGNEVIIDRIEFPVPAETFEAYDYLERAQ 621
Qy 681 AVNGLFTSKD-ALQTSVTDYQVQAANLVECLSDLYPNERKMLMDVAKEARLVLQARN 739
Db 622 AVNELFTSSNQIGLKTVDYHIDOVSNLVECLSDPECLDEKQELSEKVKHAKRLSDERN 681
Qy 740 LLQDTGNRING--ENGWTSGTIEVAGDVLFKDRSLRLTSARSIDTETTYTYLQQID 797
Db 682 LLQDPNFRGINQLDRGWRGSTDITIQQGDDVFKNYVTLTGTFD---ECYTYLYQKID 738
Qy 798 ESLKPYTRYKLGFGISSQDLLEIKLIRHRANQIVKNVDPN---LLPDVLPVNSCGGID 853
Db 739 ESKLKAYTRYQLRGYIEDSQDLLEIYLIRYNAKHETVNVPGTGLMPLSAQSFGKCGEPN 798
Qy 854 RCSEQYVDAMALANNENGENG-NMSSDSHAFSFIHDTGSEIDINENTGIWVVFPIPTNGY 912
Db 799 RCA--PHLEWNPDLDCRDCGKCAHSHHFLSDIDVGTDLNEDLGWVYFKITQDGH 856
Qy 913 ATLNLELVEGPIUSGETLERAQOQOQOMARKRGASEKAYVAAQAOIIDLPLADYQD 972
Db 857 ARLANLEFLEKPLVGEALARVRAKAKWRKREKLEMETNIVYKEAKESVDALPVNSQY 916
Qy 973 QKLSGVEMSDMLAAQNLVQIPYVYNDALPEIPGMNNTSFTLNRLOQANWLYDLRVA 1032
Db 917 DOLQADYTNAMIHAADKRVHSIREAYLPELSVPGVNAAIPEELGRIPTAFSLDARNV 976
Qy 1033 IPNGDFRNLSDMNATSDVNV-QQLSDTSVLVIPNNSQVSOQFTVQPNRYVLTATK 1091
Db 977 IKNGDFNGLSCHWNVKGVHDVEEQNHNSVLVPEWEAEVSOEVKPCRGYILRVATYK 1036
Qy 1092 EGVGQGVYIIRDGANQTTLTFN-----ICDD----- 1118
Db 1037 EGYGEGCVTHIEIKNTDDELKESNCVEEVPNNVTTCNDYATQOEYEGYTSNRGYN 1096
Qy 1119 DTGVLSDQTS-----YITKTVETPSTEQWID 1147
Db 1097 EAPSPADYASVYEKSYTDGRRNCPFRNCRGYRDTPLPVGIVTKLEYETPDKVWIE 1156
Qy 1148 MSETEGVFNIESVELVLEE 1167
Db 1157 IGETEGTFIVDSVELLME 1176

RESULT 13
JC2219Parasporal crystal protein cryIIAa - *Bacillus thuringiensis*C:Species: *Bacillus thuringiensis*

C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 01-Dec-2000

C:Accession: JC2219

R:Udayasuriyan, V.; Nakamura, A.; Mori, H.; Masaki, H.; Uozumi, T.

Biosci. Biotechnol. Biochem. 58, 830-835, 1994

A:Title: Cloning of a new cryIIA(a) gene from *Bacillus thuringiensis* strain FU-2-7 and a

A:Reference number: JC2219; MUID:94289859; PMID:7764972

A:Accession: JC2219
A:Molecule type: DNA
A:Residues: 1-1176 <UDA>
A:Cross-references: DDBJ:DI17518; NID:g506190; PIDN:BAA04468.1; PID:g535781
C:Genetics:
A:Gene: cryIA(a)
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.8%; Score 2045; DB 2; Length 1176;
Best Local Similarity 39.0%; Pred. No. 1.9e-110;
Matches 479; Conservative 174; Mismatches 415; Indels 152; Gaps 26;

QY 53 NPGL-----PCNPET-FISSSTVQTG---IGI-----VGVGLGALGVPPAGQIASFYSP 97
DB 4 NPINECIPYCNLSNPEVEVLGGERIETGTPIDISLSLTQPLSEFPVPGAGFVLGLVDI 63
QY 98 IVGQLWPSSSTVSWEMIMKQVEDLIDQKIDTSVRKTAALAGLQGLGDLDVYQKSLKNWLE 157
DB 64 IWGIFGPSQ---WDAPLVQIEQLINQRIEFPARNQAISRLGLSNLYQIYAESFREWBA 119
QY 158 NRNDTRARSVVVTVYIALELDFVAKIPSPAISGQVEPLLSVYQAANLHLLLRDASIFG 217
DB 120 DPTNPALREMRIOFNDMSALTTAIPLLAVQNVQVPLLSVYQAANLHLLSVLRDVSFVG 179
QY 218 AEWGFTPGEISTFYDRQVTTAQSYDVCVKMYNTGLDKLKTNAASMLKYHQFRREMTLL 277
DB 180 QRWGLDVATINSRYNDLTRIGTYTDAVRMYNTGLERVWGPDSRDWRYNQFRBELTIT 239
QY 278 VLDLVALFPNYDTRTYPIETTAQLTRVYTDPIVFNRETSGGCRWRWSLNSDISPSEVES 337
DB 240 VLDIVSLFPNYSRTYPIRTVSQLTRTYTPVLENFD--GSFRG-----SAQRIEQ 289
QY 338 AVTRSPHLFDILSIEFTTTRAGLPLANTTEYLEWVGHSHK-----YKNTWASSALERNY 392
DB 290 S-IRSPHMLDILNSITTYTDAHG-----GYTWSHQIMASVPVSGPBPFPPL---Y 338
QY 393 GTTISNKIKYDYLANKDIPQVRSILGADLANYAQ--VYGVVPASFTLLDKNT-----GSGS 446

913 ATLGNLLEVEGPLSGSETLERAQQEQOQWODKMKRKGASEKAYAAKQAIIDLRFADYOD 972
DB 857 ARLGNLLEFLBEPKPLVGEARLAKRKKRKKLEWETNIVYKREKESVDALFVNSQY 916
QY 973 QKLSNGVEMSDMLAAQVLQOSIPVYVNDALPEIPGMYNTSFTELTNRLQQAWNLYDLNA 1032
DB 917 DQLQADTNIAHIAADKRVHSIREAYLPVLPVPGVNAALFEELGRISTAFSLYDARNV 976
QY 1033 IPNGDFRNLGSLDNWATSDVNV--QQLSDTSVLVTPNNWSQVSQOFTVQPNRYVLRVYARK 1091
DB 977 IKNGDFNGLSCWNVKGVHDVEBQNNQSRVSLVPEWEARVSOEVRVCPGRGYILLRYAYK 1036
QY 1092 EGVGDSGVVIRDCANCTETLTEN-----ICDD----- 1118
DB 1037 EGYGEGCVTHIETENMTDELKSCVBEIYPNNTVTCNDYTVNQBEYGGAYTSRNGYN 1096
QY 1119 DTVLSADOTS-----YITKTVEFTPTPTEQVWID 1147
DB 1097 EAPSVPADYASVVEKSYTDGRRNPCEPFRNGRYDTPLPVGYVTKLELVPETDKVWIE 1156
QY 1148 MSETGVEFNIESVELVLEHE 1167
DB 1157 IGETEGTFIVDSVELLMBE 1176

RESULT 14
A41052
Parasporal crystal protein cryAel - Bacillus thuringiensis (strain alesti)
C:Species: Bacillus thuringiensis
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 01-Dec-2000
C:Accession: A41052
R:Lee, C.S.; Aronson, A.I.
J. Bacteriol. 173, 6635-6638, 1991
A:Title: Cloning and analysis of delta-endotoxin genes from Bacillus thuringiensis subsp.
A:Reference number: A41052; MUID:92011442; PMID:1655719
A:Accession: A41052
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1181 <LEB>
A:Cross-references: GB:M65252; NID:g142874; PIDN:AAA22410.1; PID:g142875
C:Superfamily: parasporal crystal protein
C:Keywords: delta-endotoxin

Query Match 33.8%; Score 2044.5; DB 2; Length 1181;
Best Local Similarity 39.0%; Pred. No. 1.9e-110;
Matches 477; Conservative 190; Mismatches 404; Indels 151; Gaps 29;

QY 53 NPGL-----PCNPET-FISSSTVQTG---IGI-----VGVGLGALGVPPAGQIASFYSP 97
DB 4 NPINECIPYCNLSNPEVEVLGGERIETGTPIDISLSLTQPLSEFPVPGAGFVLGLIDL 63
QY 98 IVGQLWPSSSTVSWEMIMKQVEDLIDQKIDTSVRKTAALAGLQGLGDLDVYQKSLKNWLE 157
DB 64 IWGIFGPSQ---WDAPLVQIEQLISQRIEFPARNQAISRLGLSNLYQIYAEPREWEA 119
QY 158 NRNDTRARSVVVTVYIALELDFVAKIPSPAISGQVEPLLSVYQAANLHLLLRDASIFG 217
DB 120 DPTNPALREMRIOFNDMSALTTAIPLLAVQNVQVPLLSVYQAANLHLLSVLRDVSFVG 179
QY 218 AEWGFTPGEISTFYDRQVTTAQSYDVCVKMYNTGLDKLKTNAASMLKYHQFRREMTLL 277
DB 180 QRWGLDVATINSRYNDLTRIGTYTDAVRMYNTGLERVWGPDSRDWRYNQFRBELTIT 239
QY 278 VLDLVALFPNYDTRTYPIETTAQLTRVYTDPIVFNRETSGGCRWRWSLNSDISPSEVES 337
DB 240 VLDIVSLFPNYSRTYPIRTVSQLTRTYTPVLENFD--GSFRG-----SAQRIEQ 289
QY 338 AVTRSPHLFDILSIEFTTTRAGLPLANTTEYLEWVGHSHK-----YKNTWASSALERNY 392
DB 290 S-IRSPHMLDILNSITTYTDAHG-----GYTWSHQIMASVPVSGPBPFPPL---Y 338
QY 393 GTTISNKIKYDYLANKDIPQVRSILGADLANYAQ--VYGVVPASFTLLDKNT-----GSGS 446

Db 899 TNIAHHAADKRVHSIREAYLPBLSVIPGUNAALFEELEGRIFTAFSLYDARNVKNKDF 958
Qy 1039 RNLSDNATSDNV-QQLSDTSVLVIPNMSQVSOQFTVQPNRYVYLVLTARKEGVGDG 1097
Db 959 NGLSCHNVKGVHVDVEQNNHRSVLVWPEAEVSVQVRVCPGRGYILRVTVAYKEGYGEG 1018
Qy 1098 YVIRDCANQTEILTFN-----ICDDDTGV----- 1122
Db 1019 CVTIHEIENNTDELKFSNCVEEYVNNVTVCNDYTATQBEYEGTTSRNRGYDAYS 1078
Qy 1123 --LSADQTS-----YITKVETPSTEQVWIDMS 1149
Db 1079 SSVPADYASAYBEKAYTDGRDNPCEBNGYDTPPLPAGYVTKELYFPETDKWIEIG 1138
Qy 1150 ETEGVFNIESVELVLEE 1167
Db 1139 ETEGTFIVDSVELLMEE 1156

Search completed: June 21, 2004, 13:39:44
Job time : 34 secs

Db 1 MSPNNQNEYIIDATPSTSVSSDSNRYPFAPEPTDALQNNYKDYLMKSGGNGENPELFCNP 60
Qy 61 ETFISSVTQGIQVIGVGLGALGVPFAGQIASFYSFIYVGLQWPSSTSVVMEIMKQVED 120
Db 61 ETFISSVTQGIQVIGVGLGALGVPFAGQIASFYSFIYVGLQWPSSTSVVMEIMKQVED 120
Qy 121 LIDOKITDVKRTALAGLOGDGLDVTQKSLKWLNRNDTRARSVVVYIYALDLDFV 180
Db 121 LVDOKIKRYKDKALAEKGLGNALDVTQKSLKWLNRNDTRARSVVVYIYALDLDFV 180
Qy 181 AKIPSAISGEVPLLSVYAQAANLHLLLDASIFGAEMGFTPEISSTFYVDROVTTAQ 240
Db 181 SSISFVAVSGHEVLLAVYAQVNLHLLLDASIFGEEMGFTPEISSTFYVDROVTTAQ 240
Qy 241 YSDYCVWYNTGLDKLGTNAASLKYHQFPRMTLLVLDLVALFPNPDYRTYETTTAQ 300
Db 241 YSDYCVWYKGLDKLGTTSKMLNYHQFPRMTLLVLDLVALFPNPDYRTYETTTAQ 300
Qy 301 LTRVYVTDPIVNRETSGGFCRRVSLNSDISFSEVESAVIRSPHLFDILSIEFTYTRAG 360
Db 301 LTRVYVTDPIAFNIVTSTGFCNPWSTHSGILFYEVNNVIRPPELFDILSIEFTYTRAG 360
Qy 361 LPLANTYEVWVCHSFKYNT-NASSALERNYGTITSNKIKYYDLANKDIFOVRSIGAD 419
Db 361 ITLANDAYINWGHTRKRTASTVITYTANYGRITSEK-NSPALEDORDIFEINSTVAN 419
Qy 420 LANYAQQVYGVYFASFTLLDKNTGSGVGGFTYSKPHYTMQVCTQNTYDIBEPENE-P 478
Db 420 LANYQKAYGVGSFHWKX-CTSSITATLYSKTHALOGCTQVYESSDEILDRTPV 477
Qy 479 LSRGYSHRLSHITSYSPSKNASSPARYGNLVPFANTHRASDVNTVYSDDKLTQIPVVKAH 538
Db 478 VABSYSHRLSHITSYSPSKNASSPARYGNLVPFANTHRASDVNTVYSDDKLTQIPVVKAH 535
Qy 539 TLVSGTTLVKGPGTGNILARTSGGLAYTSVSKPLSQRYPARYASTNLRFLVFT 598
Db 536 MLXGGSVVQPGFTGGDILKRTPSILGTFAVTVNGSLQRYRIRKRYASTNLRFLVFT 594
Qy 599 ISGTRIYSINVKTMKDDTLFTNFDLATIGTAFYTFNSYDSLTIVGADSPASGEVYVD 658
Db 595 YLGDITBKRPNKTMONGASLYTETFPKASFTIDPQFRETQDKILLSGDFSSGQEVYID 654
Qy 659 KPELIPVNTAEABEDLDVAKAVNGLETSKDALOTSVDYQVNOANLVECLSDLYP 718
Db 655 RIEFIPVDYEAQDLAEAKKAVNALFTNTKGLRPGVTDYEVNOANLVECLSDLYP 714
Qy 719 NEKMLMDAVKAKLVQARNLQDTPGNRNGENGWGTGIEVAEGDVLFPDRLSLT 778
Db 715 NEKMLMDAVKAKLVQARNLQDTPGNRNGENGWGTGIEVAEGDVLFPDRLSLT 774
Qy 779 SAREIDTETPTLYQQIDELLPYTRYKLGKFGSSQDLLEIKLIRHRANQIVKNVPDN 838
Db 775 GAREIDTETPTLYQVKEGVLPYTRYLRGFGVSSQGLEIVTIRHQTNRVKNVPPDD 834
Qy 839 LLPVLVWNSCGGIDRCSEQQYVDANLALNNGENGWSSSHAPSHFDGCEIDLNENT 898
Db 835 LLPVSPVNSDGSINRCSEQKYNSRL-----EGENRSGDAHEFSLPIDIGELDYNENA 888
Qy 899 GIWVVPKLPITNGVATLGNLELVEBPLSGETLERAQOQOQWQDKMARKEGASKAYYA 958
Db 889 GIWVGFKLTPDEGATLGNLELVEBPLSGDALERQEBQOMKQTRRREEDDRRTMA 948
Qy 959 AKQADRLFADYQDQKLSGVEMSDMLAAQNLVQSIPIVYNDALPEIPGMNYTSTFELTN 1018
Db 949 SKQAVDLRYADYQDQNLDPVEITDLTAAQDLIQSIPVYVNEPPEIPGMNYTSTFELTD 1008
Qy 1019 RLQQAANLYDLRNAIPNGDFRNLGSDNATSDNVVQQLSDTSVLVIPNWSQVSOQFTVQ 1078
Db 1009 RLQQAANLYDQRNAIPNGDFRNLGSDNATSDNVVQQLSDTSVLVIPNWSQVSOQFTVQ 1069
Qy 1079 PNRYVVLVATKKGVGQYVIRIGANOTETLTENICDDDTGVLSADQTSVITKTVFT 1138
Db 1069 PNQRYVVLVATKKGVGQYVIRIGANOTETLTENICDDDTGVLSADQTSVITKTVFT 1128

Qy 1139 PSTQVWIDMSSETGVFNIESVELVLBER 1167
Db 1129 PYTDQWIEISETGTFVIESVELIVDVE 1157

RESULT 2

C8CA BACTP STANDARD; PRT; 1160 AA.
ID C8CA BACTP AC Q45706;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pesticidal crystal protein crysca (insecticidal delta-endotoxin
DE CRYVIIIC(a) (Crystalline entomocidal protoxin) (130 kDa crystal
DE protein).
GN CRY8CA OR CRYVIIIC(A) OR CRYIIII.
OS Bacillus thuringiensis (subsp. japonensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OK NCBI_TaxID=128936;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Buibui.
RX MEDLINE=94100786; PubMed=7764305;
RA Sato R., Takeuchi K., Ogiwara K., Minami M., Kaji Y., Suzuki N.,
Hori H., Asano S., Ohba M., Iwahana H.;
RT "Cloning, heterologous expression, and localization of a novel crystal
protein gene from Bacillus thuringiensis serovar japonensis strain
bulbui toxic to scarabaeid insects.";
RL Curr. Microbiol. 28:15-19(1994).
RN (2)
RP SEQUENCE OF 1-14 AND 56-64, AND CHARACTERIZATION.
RX MEDLINE=94259659; PubMed=820856;
RA Hori H., Suzuki N., Ogiwara K., Himejima M., Indrasith L.S.,
Minami M., Asano S., Sato R., Ohba M., Iwahana H.;
RT "Characterization of larvicidal toxin protein from Bacillus
thuringiensis serovar japonensis strain Buibui specific for
scarabaeid beetles.";
RL J. Appl. Bacteriol. 76:307-313(1994).
CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
EPITHELIAL CELLS OF INSECTS. ACTIVE ON VARIOUS SCARABAEID BEETLES
SUCH AS ANOMALA CUPREA, A. RUFOCUPREA AND POPILLIA JAPONICA.
CC -1- DEVELOPMENTAL STAGE: The crystal protein is produced during
sporulation and is accumulated both as an inclusion and as part of
the spore coat.
CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-
terminus.
CC -1- SIMILARITY: Belongs to the delta endotoxin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; U04366; AAA21119.1; -
DR PIR; I40589; I40589.
DR HSP; P07130; 1DLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
KW Toxin; Sporulation.
SQ SEQUENCE 1160 AA; 130425 MW; C16C3D912EBB8751 CRC64;

Query Match 61.8%; Score 3733.5; DB 1; Length 1160;
Best Local Similarity 64.2%; Pred. No. 4.4e-205;
Matches 767; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

```
QY 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKYDYLMSSEGENPELPGNP 60
DB 1 MSPNNQNEYIILDASSPTSVSDNSRYPLANDQTTTLQNNYKYDYLKWTSTNAELSRNP 60
QY 61 ETPFIS-STVQTGIGVGLGALGVPFAGQIASFYSPVIGOLWPSSTVSVMWEMKQVE 119
DB 61 GTFISAQDAVGTDIVSTIISGLGIPVLGEVFSILGSLIGLLWFSNNENVMQIIPNVRV 120
QY 120 DLIDQKITDSVRKTAALAGLQGLGDVYQKSLKNWLENNDTRARSVVVTVQYIALELDF 179
DB 121 ELIDOKIILDSVRRAIADLANSLRIAVEYQNALEDWRKNPHSTRNALVKERPGHAIL 180
QY 180 VAKIPSPAISSQGVFLSVYAQAANLHLLLRDASIFGAEWGFTPGEISTFYDROVTKTA 239
DB 181 RTNMGSPSTNYETETPLPTAQAASLHLVMDVQVIGKEMWGPQNDIDLKYEQVSYTA 240
QY 240 QYSDYCVKQWYNTGLDKLGTNAASHLKYHOPREMLTLLVLAVALPNNVDTETPIETA 299
DB 241 RYSDHCQWYNAGLNKLGRTGAKQWVDYRFRFNNVMVLDLVALFPNDARIYPLETNA 300
QY 300 OLTRVYTDPIVFNRETSQGFCKRWSLN-----SDI-----SFSEVESAVIRSPHELDIL 349
DB 301 ELTREIFIDPV-----GSYVTGSSSTLISYWDMPAALPSPSTLEN-LARKPDFFTL 352
QY 350 SBIETPTTRAGLPLNNT-EYLEYVWGHSLIKYKNTNASSALERNYGTITSNKIKYYDLANK 408
DB 353 QBIRMYTS---PRQMTIHYNNWGGQRLTSLYVGGSP--NKYSGLVLAGAEDIIIPVQGN 407
QY 409 DIPQVRSIGADLANTYAQVGYVPYASFLLDIDWTCGSGVGGTYSPKPTMTQVCTQNTY 468
DB 408 DIYRV--VMTYIGRTYNSLLGNVPYTF-YFSNNYTK-----TYSKP-----KOFAGGIKT 454
QY 469 ID---EIPPENEPLSGYSHRLSHITSYFSKNASPARYGNLPPAMTHRSADVTNTVY 525
DB 455 IDSGBELTYEN---VQSISRHSYVITSFEIKSTGTV--LGVVPFGWTHSSASRNNFY 509
QY 526 SDKITQIPVVKALVSGTIVTKGPG-FTGQNLKRTSGGLAY---TSVSVKSPLSQRY 581
DB 510 ATKISQIPINKASRTSGGAVNPFQGLYNGPVMKLSGSGSVINLRVATDAKG-ASQRY 568
QY 582 RARIYASTNLRFLWTIS-----GTFIYINVNKTANKGDDLTFTNFDLATIGTA-F 633
DB 569 RIRIYASDRAGKP--TISSRSPENPATYSASIAVTNTMTNASTYSTFAVASEGPIML 626
QY 634 TFSNYSDSLTVGADSPASGGEVYVDKFLIPVNATFEAEDLDVAKAVNGLFTSKDAL 693
DB 627 GISGSRPTDISITKBAGNANIYDRIEPIPVNTLFEAEDLDVAKAVNGLFTKEKDAL 686
QY 694 QTSVTDYQVNOAANLVECLSDLYPNKRMMLMDAVKBAKRLVOARNLLQDTGFNRINGEN 753
DB 687 QTSVTDYQVNOAANLVECLSDLYPNKRMMLMDAVKBAKRLVOARNLLQDTGFNRINGEN 746
QY 754 GWTGTGTGEVAGDVLFKDRSLRLSAREIDTETPTLYLYQOIDSLLKPYTRYKLGFI 813
DB 747 GWTGTGTGEVAGDVLFKDRSLRLSAREIDTETPTLYLYQOIDSLLKPYTRYKLGFI 806
QY 814 GSSQDLKILIRHRANOIVKNVPMNLLPDVLPVNSCGGIDRCSQOYVDANLALENNGEN 873
DB 807 GSSQDLKILIRHRANOIVKNVPMNLLPDVLPVNSCGGVDRCSQOYVDANLALENNGEN 866
QY 874 GNMSSDSHAFSPHIDTGEIDLNTGNTGIWVVFVKIPPTNGYATIGNLELVEBGPLSGETLER 933
DB 867 GNMSSDSHAFSPHIDTGEIDLNTGNTGIWVVFVKIPPTNGYATIGNLELVEBGPLSGETLER 926
QY 934 AQOQSQOWDKWARKGASEKAYAAKQAI DRLPADYQOKLNSGVMSDMLAAQNLVQS 993
DB 927 AQOQSQOWDKWARKGASEKAYAAKQAI DRLPADYQOKLNSGVMSDMLAAQNLVQS 986
QY 994 IPVYNDALPEIPGMNYTSPTELNRLOQAWNLXDLRNAL PNGDPFNGLSDMNATSDVNV 1053
DB 987 IPVYNDALPEIPGMNYTSPTELNRLOQAWNLXDLRNAL PNGDPFNGLSNMNATSDVNV 1046

QY 1054 QQLSDTSVLVPIPNWNSQVSOQTVQPNRYVLRVARTKSGVGDGVYVLRDGNQOTETLTF 1113
DB 1047 QQLSDTSVLVPIPNWNSQVSOQTVQPNRYVLRVARTKSGVGDGVYVLRDGNQOTETLTF 1106
QY 1114 NICDDDTGVLSDAQTSYITKTVFTPTSTBQWMDMSBTEGVFNIESVELVLEE 1167
DB 1107 NICDDDTGVLSDAQTSYITKTVFTPTSTBQWMDMSBTEGVFNIESVELVLEE 1160

RESULT 3
C8BA BACUK
ID C8BA BACUK STANDARD; PRT; 1169 AA.
AC Q45705;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cry8Ba (Insecticidal delta-endotoxin
DE CryVIIIa(a)) (Crystalline entomocidal protoxin) (134 kDa crystal
DE protein).
GN CRY8BA OR CRYVIII(A) OR 50C(B).
OS Bacillus thuringiensis (subsp. kumamotoensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=132267;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NRRL B-18746 / PS50C;
RA Michaelis T.B., Fonceirada L., Narva K.E.;
RT "Process for controlling scarab pests with Bacillus thuringiensis
RT isolates";
RL Patent number WO9315206, 05-AUG-1993.
CC -I- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS. ACTIVE ON VARIOUS SCARABID BEETLES.
CC -I- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -I- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -I- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U04365; AAA21118.1; --
CC HSP: P07130; LDLC.
CC InterPro: IPR001178; Endotoxin.
CC InterPro: IPR005638; endotoxin_C.
CC InterPro: IPR005639; endotoxin_N.
CC InterPro: IPR008979; Gal_bind_like.
CC Pfam: PF00555; endotoxin_1.
CC Pfam: PF03944; endotoxin_C; 1.
CC Pfam: PF03945; endotoxin_N; 1.
CC Toxin; Sporulation.
CC KW SEQUENCE 1169 AA; 133543 MW; 228BFCF5BD699909 CRC64;
SQ
Query Match 52.2%; Score 3152; DB 1; Length 1169;
Best Local Similarity 54.0%; Pred. No. 6.4e-172; Indels 84; Gaps 25;
Matches 653; Conservative 179; Mismatches 294;
QY 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKYDYLMSSEGENPELPGNP 60
DB 1 MSPNNQNEYIILDASSPTSVSDNSRYPLANDQTTTLQNNYKYDYLKWTSTNAELSRNP 60
QY 61 ETPFIS-STVQTGIGVGLGALGVPFAGQIASFYSPVIGOLWPSSTVSVMWEMKQVE 119
DB 61 GTFISAQDAVGTDIVSTIISGLGIPVLGEVFSILGSLIGLLWFSNNENVMQIIPNVRV 120
QY 120 DLIDQKITDSVRKTAALAGLQGLGDVYQKSLKNWLENNDTRARSVVVTVQYIALELDF 179
DB 121 ELIDOKIILDSVRRAIADLANSLRIAVEYQNALEDWRKNPHSTRNALVKERPGHAIL 180
QY 180 VAKIPSPAISSQGVFLSVYAQAANLHLLLRDASIFGAEWGFTPGEISTFYDROVTKTA 239
DB 181 RTNMGSPSTNYETETPLPTAQAASLHLVMDVQVIGKEMWGPQNDIDLKYEQVSYTA 240
QY 240 QYSDYCVKQWYNTGLDKLGTNAASHLKYHOPREMLTLLVLAVALPNNVDTETPIETA 299
DB 241 RYSDHCQWYNAGLNKLGRTGAKQWVDYRFRFNNVMVLDLVALFPNDARIYPLETNA 300
QY 300 OLTRVYTDPIVFNRETSQGFCKRWSLN-----SDI-----SFSEVESAVIRSPHELDIL 349
DB 301 ELTREIFIDPV-----GSYVTGSSSTLISYWDMPAALPSPSTLEN-LARKPDFFTL 352
QY 350 SBIETPTTRAGLPLNNT-EYLEYVWGHSLIKYKNTNASSALERNYGTITSNKIKYYDLANK 408
DB 353 QBIRMYTS---PRQMTIHYNNWGGQRLTSLYVGGSP--NKYSGLVLAGAEDIIIPVQGN 407
QY 409 DIPQVRSIGADLANTYAQVGYVPYASFLLDIDWTCGSGVGGTYSPKPTMTQVCTQNTY 468
DB 408 DIYRV--VMTYIGRTYNSLLGNVPYTF-YFSNNYTK-----TYSKP-----KOFAGGIKT 454
QY 469 ID---EIPPENEPLSGYSHRLSHITSYFSKNASPARYGNLPPAMTHRSADVTNTVY 525
DB 455 IDSGBELTYEN---VQSISRHSYVITSFEIKSTGTV--LGVVPFGWTHSSASRNNFY 509
QY 526 SDKITQIPVVKALVSGTIVTKGPG-FTGQNLKRTSGGLAY---TSVSVKSPLSQRY 581
DB 510 ATKISQIPINKASRTSGGAVNPFQGLYNGPVMKLSGSGSVINLRVATDAKG-ASQRY 568
QY 582 RARIYASTNLRFLWTIS-----GTFIYINVNKTANKGDDLTFTNFDLATIGTA-F 633
DB 569 RIRIYASDRAGKP--TISSRSPENPATYSASIAVTNTMTNASTYSTFAVASEGPIML 626
QY 634 TFSNYSDSLTVGADSPASGGEVYVDKFLIPVNATFEAEDLDVAKAVNGLFTSKDAL 693
DB 627 GISGSRPTDISITKBAGNANIYDRIEPIPVNTLFEAEDLDVAKAVNGLFTKEKDAL 686
QY 694 QTSVTDYQVNOAANLVECLSDLYPNKRMMLMDAVKBAKRLVOARNLLQDTGFNRINGEN 753
DB 687 QTSVTDYQVNOAANLVECLSDLYPNKRMMLMDAVKBAKRLVOARNLLQDTGFNRINGEN 746
QY 754 GWTGTGTGEVAGDVLFKDRSLRLSAREIDTETPTLYLYQOIDSLLKPYTRYKLGFI 813
DB 747 GWTGTGTGEVAGDVLFKDRSLRLSAREIDTETPTLYLYQOIDSLLKPYTRYKLGFI 806
QY 814 GSSQDLKILIRHRANOIVKNVPMNLLPDVLPVNSCGGIDRCSQOYVDANLALENNGEN 873
DB 807 GSSQDLKILIRHRANOIVKNVPMNLLPDVLPVNSCGGVDRCSQOYVDANLALENNGEN 866
QY 874 GNMSSDSHAFSPHIDTGEIDLNTGNTGIWVVFVKIPPTNGYATIGNLELVEBGPLSGETLER 933
DB 867 GNMSSDSHAFSPHIDTGEIDLNTGNTGIWVVFVKIPPTNGYATIGNLELVEBGPLSGETLER 926
QY 934 AQOQSQOWDKWARKGASEKAYAAKQAI DRLPADYQOKLNSGVMSDMLAAQNLVQS 993
DB 927 AQOQSQOWDKWARKGASEKAYAAKQAI DRLPADYQOKLNSGVMSDMLAAQNLVQS 986
QY 994 IPVYNDALPEIPGMNYTSPTELNRLOQAWNLXDLRNAL PNGDPFNGLSDMNATSDVNV 1053
DB 987 IPVYNDALPEIPGMNYTSPTELNRLOQAWNLXDLRNAL PNGDPFNGLSNMNATSDVNV 1046
```

Db 121 ELINQIAEYARNKALSELBLEGNYYQLYLTALBEMKENPNSRALDVRNRFILDSLF 180
Qy 180 VAKIPSPAGSQORVPLSVYAQAANLHLLLRDASIFGAEWGFTPGISTFYDRQVTRTA 239
Db 181 TOYMPSPRVNPFVFFLTVTYMAANLHLLLRDASIFGEWGLSTSTINNYNQMKLTA 240
Qy 240 QYSDYCVKMYNTGLDKLGTNAASMLKXHOFRREMTLLVLDLVALPNYDTRTYPIETTA 299
Db 241 EYSDHCWKVYETGLAKLGSQAKQWIDYNQRRMTLLVLDLVALPNYDTRTYPIETTA 300
Qy 300 QLTREVTDPVFNKRETSGGCRWLSNDSISFSEVASAVIRSPLHLDLSEIEFYFTRA 359
Db 301 QLTREVTDPVFNKRETSGGCRWLSNDSISFSEVASAVIRSPLHLDLSEIEFYFTRA 356
Qy 360 GLPLNTEYLEYVGHSHIKYKQNTNASSALERNYGT-----ITSNKIKYDLANKDIFQVR 414
Db 361 GLPLNTEYLEYVGHSHIKYKQNTNASSALERNYGT-----ITSNKIKYDLANKDIFQVR 414
Qy 357 S--FTSDRYERYWAGHQISYKHGISTFTQMYGTQNLQSTN----FDPNTYDIYKTL 410
Db 358 S--FTSDRYERYWAGHQISYKHGISTFTQMYGTQNLQSTN----FDPNTYDIYKTL 410
Qy 415 SLGADLAN-----YAAQVGVYASFTLLDKNTGSGVGGFTYSKPHHTMQVCTQNTYI 469
Db 416 SLGADLAN-----YAAQVGVYASFTLLDKNTGSGVGGFTYSKPHHTMQVCTQNTYI 469
Qy 411 SNGAVLLDIVPGYTYTFFGMPETEPFPMVQNLNTRKT--LTY-KP-ASKDIIDTRDSE 466
Db 412 SNGAVLLDIVPGYTYTFFGMPETEPFPMVQNLNTRKT--LTY-KP-ASKDIIDTRDSE 466
Qy 470 DEIPPE--NEPLSGYSHRLSHITSYFSKNASPARYGNLPVPAWTHRSADVTNVTYSD 527
Db 471 DEIPPE--NEPLSGYSHRLSHITSYFSKNASPARYGNLPVPAWTHRSADVTNVTYSD 527
Qy 467 LELPETSQGNYESYSHRLGHIT-FYSSSTST-----YVPVSWTHRSADLTNTVWSG 520
Db 468 LELPETSQGNYESYSHRLGHIT-FYSSSTST-----YVPVSWTHRSADLTNTVWSG 520
Qy 528 KITQIPVKAHTLVSGTIVKPGFTGNNILKRTSSGPLAYTSVSKSPLSQRYRIRY 587
Db 529 KITQIPVKAHTLVSGTIVKPGFTGNNILKRTSSGPLAYTSVSKSPLSQRYRIRY 587
Qy 521 BITQIPGKSSTIGENTVYIKRGVGTGDLVALTDR--IGSCBQOMIPFESQRFIRY 578
Db 522 BITQIPGKSSTIGENTVYIKRGVGTGDLVALTDR--IGSCBQOMIPFESQRFIRY 578
Qy 588 AS--TTNRLP-FYISGTRIYSINVKT-MKQGD-DLTFNTPDLATGTATFNSYDSL 642
Db 589 AS--TTNRLP-FYISGTRIYSINVKT-MKQGD-DLTFNTPDLATGTATFNSYDSL 642
Qy 579 ASNETSYISGLNQSGT-----LKFQNTYSKNENDLYTND-----PKYIEYPRVI 625
Db 580 ASNETSYISGLNQSGT-----LKFQNTYSKNENDLYTND-----PKYIEYPRVI 625
Qy 643 TVGADS-----PASGEVYV-DKPELIPVNATERAEEDLDVAKAVNGLFTSKDAL 693
Db 644 TVGADS-----PASGEVYV-DKPELIPVNATERAEEDLDVAKAVNGLFTSKDAL 693
Qy 626 SVNASSIQRISIGIQNTNLFILDRIFIPDETYAEATDLEAKAVNGLFTSKDAL 685
Db 627 SVNASSIQRISIGIQNTNLFILDRIFIPDETYAEATDLEAKAVNGLFTSKDAL 685
Qy 694 QTSYTDYQVNOANLVECLSDLYPNKRMLEMDAWEAKRLVQARNLLQDTGFRNNGEN 753
Db 695 QTSYTDYQVNOANLVECLSDLYPNKRMLEMDAWEAKRLVQARNLLQDTGFRNNGEN 753
Qy 686 QPGVTDYEVNOANLVECLSDLYPNKRMLEMDAWEAKRLVQARNLLQDTGFRNNGEN 745
Db 687 QPGVTDYEVNOANLVECLSDLYPNKRMLEMDAWEAKRLVQARNLLQDTGFRNNGEN 745
Qy 754 GWTSGTGLEAVGDLVKORSLRLTSAREIDTETPTLYLYQOIESLLKPYTRYKLGPI 813
Db 755 GWTSGTGLEAVGDLVKORSLRLTSAREIDTETPTLYLYQOIESLLKPYTRYKLGPI 813
Qy 746 GWTASTGIEVTEGDAVFGKRYLRLFGAREIDTETPTLYLYQKVEGVLKPYTRYLRGFV 805
Db 747 GWTASTGIEVTEGDAVFGKRYLRLFGAREIDTETPTLYLYQKVEGVLKPYTRYLRGFV 805
Qy 814 GSSODLEKLIRHANOIVKVPNLLPDVLPVNSCGGIDRCSCQVYDAMLLENNGEN 873
Db 815 GSSODLEKLIRHANOIVKVPNLLPDVLPVNSCGGIDRCSCQVYDAMLLENNGEN 873
Qy 806 GSSGLEIYTHQTNRLVKNVPDOLLDPVFPVNDGRINRCSEKQVNSRLEVENR--- 862
Db 807 GSSGLEIYTHQTNRLVKNVPDOLLDPVFPVNDGRINRCSEKQVNSRLEVENR--- 862
Qy 874 GNMSSDSHAFSHIDTGBIDNENTGIWVVFKEIPTNGYATLGNLELVEEGPLSGETLER 933
Db 875 GNMSSDSHAFSHIDTGBIDNENTGIWVVFKEIPTNGYATLGNLELVEEGPLSGETLER 933
Qy 863 ---SGBAEHFSIPIDTGBELDENAGIWWGPKITDPEGYATLGNLELVEEGPLSGDALER 919
Db 864 ---SGBAEHFSIPIDTGBELDENAGIWWGPKITDPEGYATLGNLELVEEGPLSGDALER 919
Qy 934 AQOQEQOQWQARKGASAKYAAQAIIDRLFADYQDOKLNSGVMSDMLAQLVQS 993
Db 935 AQOQEQOQWQARKGASAKYAAQAIIDRLFADYQDOKLNSGVMSDMLAQLVQS 993
Qy 920 LQKEEQWQKIQWTRERRETDRYMAKQAVDLYADYQDQNLNPNVEITDLTAQDLQS 979
Db 921 LQKEEQWQKIQWTRERRETDRYMAKQAVDLYADYQDQNLNPNVEITDLTAQDLQS 979
Qy 994 IPYVNDALPEIPGMANTYSFTLNTLQOANLVDLRNAIENGDFRNLSDNNATSDVNV 1053
Db 995 IPYVNDALPEIPGMANTYSFTLNTLQOANLVDLRNAIENGDFRNLSDNNATSDVNV 1053
Qy 980 IPYVNEFPPIPGNYTKFTLDRLOQAGLVDORNAIENGDFRNLSDNNATSGVNV 1039
Db 981 IPYVNEFPPIPGNYTKFTLDRLOQAGLVDORNAIENGDFRNLSDNNATSGVNV 1039
Qy 1054 QQLSDTSVLVLPNNNSQVSOQFTVPQNYRVYVLRVTKRKGVDGQVYIRDGANQETLTF 1113
Db 1055 QQLSDTSVLVLPNNNSQVSOQFTVPQNYRVYVLRVTKRKGVDGQVYIRDGANQETLTF 1113
Qy 1040 QQINHTSVLVLPNNNEQVSOQFTVPQNYRVYVLRVTKRKGVDGQVYIRDGANQETLTF 1099
Db 1041 QQINHTSVLVLPNNNEQVSOQFTVPQNYRVYVLRVTKRKGVDGQVYIRDGANQETLTF 1099
Qy 1114 NICDDTGLVSADQTS-----YIT-----KTFEPTPSTEQWIDHSETGEPNI 1157
Db 1115 NICDDTGLVSADQTS-----YIT-----KTFEPTPSTEQWIDHSETGEPNI 1157
Qy 1100 SASDYDTNGMYDTQASNTNGNTNTSVVMYKPAISGRKTVDISVSNQNMWIEISSETGYFI 1159
Db 1101 SASDYDTNGMYDTQASNTNGNTNTSVVMYKPAISGRKTVDISVSNQNMWIEISSETGYFI 1159
Qy 1158 ESVELVLREE 1167
Db 1159 ESVELVLREE 1167
Qy 1160 ESVELLVDFE 1169
Db 1161 ESVELLVDFE 1169

RESULT 4

C9CA_BACTO STANDARD; PRT; 1157 AA.
ID_C9CA_BACTO
AC 045733;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DB Pesticidal crystal protein cryCa (insecticidal delta-endotoxin
CryIXC(a)) (Crystalline entomocidal protoxin) (130 kDa crystal
protein).
DB CRY9CA OR CRYIXC(A).
GN Bacillus thuringiensis (subsp. tolworthi).
OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=1442;
OX [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP STRAIN=BT02618A; PubMed=8572715;
RX MEDLINE=96141404; PubMed=8572715;
RA Lambert B., Buyse L., Decock C., Janssens S., Pien S., Saey B.,
Seuring J., Van Audenhove K., Van Rie J., Van Vliet A., Peferoen M.;
"A Bacillus thuringiensis insecticidal crystal protein with a high
activity against members of the family Noctuidae.";
RL Appl. Environ. Microbiol. 62:80-86(1996).
CC -|- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
EPITHELIAL CELLS OF LEPIDOPTERA LARVAE. HAS A FAIRLY BROAD
SPECTRUM OF ACTIVITY AGAINST MEMBERS OF THE FAMILIES PYRALIDAE,
PLUTELLIDAE, SPHINGIDAE, AND NOCTUIDAE. IT IS THE FIRST
INSECTICIDAL CRYSTAL PROTEIN WITH ACTIVITY AGAINST CUTWORMS. NO
ACTIVITY IS OBSERVED AGAINST SOME BEETLES, SUCH AS COLORADO POTATO
BEETLE.
CC -|- DEVELOPMENTAL STAGE: The crystal protein is produced during
sporulation and is accumulated both as an inclusion and as part of
the spore coat.
CC -|- MISCELLANEOUS: Toxic segment of the protein is located in the N-
terminus.
CC -|- SIMILARITY: Belongs to the delta endotoxin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC EMBL; Z37527; CA85764.1; -;
DR PIR; A59350; S49247.
DR HSBP; P07130; 1DLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin_C_1.
DR Pfam; PF03945; endotoxin_N_1.
KW Toxin; Sporulation.
SQ SEQUENCE 1157 AA; 129775 MW; C364391EF7DFB8A CRC64;
Query Match 51.7%; Score 3123; DB 1; Length 1157;
Best Local Similarity 52.3%; Pred. No. 2.8e-170;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;
Qy 1 MSPNNQNEYELDASSSTSVSDNSVRYPLANDQTTTLQNNYKYLRMSGEENPELFGNP 60
Db 1 MNRNNQNEYELIDAPHCCGSPDDVRYPLASDPNAAALQNNYKYLTQTDYDTSYINP 60
Qy 61 ETPFSS-STVGTGIGVQVLGALGVPPAGQIASFYISFVQGLMPSSTSVSWEMTKQVE 119
Db 61 SLSISGRDAVQTALTVGRILGALGVPPSGQIVSGYQFLNTLTLPVNDTAWEAFMRQVE 120
Qy 120 DLIDQKITDTSVRKTAALAGLQGLDGLDVTYQKSLKNWLENRNDTRARSVVVQYIAELDP 179

```
Db 121 ELVNOQITEFARNQALRLQGLGDSFNVYQSRSLQNLADENDTRNLNVRAQPIALDLDF 180
Qy 180 VAKIPSPATSGQVPLLSVYAQANLHLLLRDASIFGAEWGFTGCEISTEYDQVTRTA 239
Db 181 VNAIPFVANGQVPLSVYAQANLHLLLRDASIFGAEWGFTGCEISTEYDQVTRTA 240
Qy 240 QYSDYCVKRWYNTGLDKGTNAASWLKTHQPREMTLLVLDLVALPNNYDTRTYPIETTA 299
Db 241 KVTNYCETWYNTGLDLRGNTESWLRYHQPREMTLLVLDLVALPNNYDTRTYPIETTA 300
Qy 300 QLTREVTYDPIVENRTSGGFCRRWSLNSDISFSEVESAVIRSPHLPDILSEIETPTTRA 359
Db 301 QLTREVTYDPIVENPPANVGLCRWGTNPYNTSELENAPFIRPHLPDILSEIETPTTRA 359
Qy 360 GLPLANNTEYLEVWGHVSIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFVRSRL 416
Db 360 -FPV-SSNFMWYSGHTLRSYLNDSAVQEDSGLTITRATNPGVDGTRN-----IEST 413
Qy 417 GADLANYYAQVGVYASFTL--LDKNTGSGVGGFTYSKPHHTMQVCTQNTYNTIDEIPP 474
Db 414 AVDPFRSALIGYGNRASFPVGLFNGTTPANGG-----CRDLYTNDLBP 461
Qy 475 KNEPLSRGYS-HRLSHITSFSGKN-ASSPARYGNLPVPAWTHRSADVTVTVYSKTIQI 532
Db 462 DE---STGSSYHRLSHVTPFPQTNQAGSIANAGSVPTVYVTRDVLNNTITENRITOL 518
Qy 533 PVVKANTLVSGTIVKPGFTGGMILKRTSGGLAVTSVSKPSLQSRVRYARRYASTTN 592
Db 519 PLVKAAPVSGTIVKPGFTGGMILKRTSGGLAVTSVSKPSLQSRVRYARRYASTTN 592
Qy 593 LRLFTVISTGRIYSINVKTMKNGDDLTENTF---DLATIG---TAFPSNYSLSLTGVA 646
Db 579 FSIKVLGGVSIGDVLKSGTMRGQBLTYESFPTRFTTTGPPNPPPTPTQAQELTVNA 638
Qy 647 DSPASGGEVYVDFELIPVNAPEAREBEDLVAKAVNGLFTSKKDALQTSVTVYQVQAA 706
Db 639 EGVSTGGEYVIDRIEIVPNPAREAREBEDLVAKAVNGLFTSKKDALQTSVTVYQVQAA 698
Qy 707 NLVECLSDLEYPNKEMLDVAKKRLVQARNLQDTPGNRING--ENGMWSTGTEVA 764
Db 699 NLVSCLSDEQYGHQKMLLBAVRAAKLSRERLLQDPDFTNTSTEENGWKAANGVTIS 758
Qy 765 EGVLFKDRSLRLTSAREIDTETPTLYXQIDESLLKPYRYLKLKPGFIGSSQDLKLI 824
Db 759 EGGPFKGRALQASAR---ENYPTIYQKVDASVLKPYRYLKLKPGFIGSSQDLKLI 814
Qy 825 RHRANOIVKVPNNLIPDVLVFNSSCGIDRCSEQOYVDANLALNENGENGMSDSEHAPS 884
Db 815 HHHKVLHVKVNPDLVSDTYSKSGCSGINRCDEQHVDMLDABHHHPMDCBAAQTSEFS 874
Qy 885 FHIDTGEIDNENTGIWVVPKIPPTNGYATLGNLELVEBGLSGETLERAQOQOQWODK 944
Db 875 SYINTGDLNASVDQGIWVVLKVRTTDDGATLGNLELVEBGLSGETLERAQOQOQWODK 934
Qy 945 MARKRGASKEYAQAQIDRLFADYQDQKLNSEVMSDMLAALNLYQSIPTVYNDALPE 1004
Db 935 LGRKRAEIDRVYLAQAQINELFDYQDQKLNSEVMSDMLAALNLYQSIPTVYNDALPE 994
Qy 1005 IPGANYTSTFELNRLQAQANLIDLNAIPNGFPRNGLSDWNATSDVNVQQLSDTSVLVI 1064
Db 995 IPGANYTSTFELNRLQAQANLIDLNAIPNGFPRNGLSDWNATSDVNVQQLSDTSVLVI 1054
Qy 1065 PNWNSQVSOQFTQVPPYRYVLTARKEGVDGVIIRDCGANQOTETLTFCDDDTGVL 1124
Db 1055 SHWDAQVSOQLRVNPNCKYLVRTARKVGGDGVYVIRDCGANQOTETLTFCDDDTGVL 1114
Qy 1125 ADQTSYTKTVEPTFTEQVWIDMSSETGVFNIESVELVLEE 1167
Db 1115 VNDNSYITEEVFPYPTKHMVVESESGSFYDISEFIETQE 1157
```

RESULT 5

```
C9EA_BACTA STANDARD; PRT; 1150 AA.
AC Q9ZNL9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cry3a (insecticidal delta-endotoxin
DE CRYIXE(a)) (Crystalline entomocidal protoxin) (130 kDa crystal
DE protein).
GN CRY9EA OR CRYIXE(A).
OS Bacillus thuringiensis (subsp. aizawai).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1433;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSK-10;
RA Midoh N., Oyama K.;
RT "Bacillus thuringiensis cry gene for insecticidal crystal protein.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -I- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -I- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -I- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB011496; BAA34908.1; -.
DR HSSP; P07130; IDLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
KW Toxin; Sporulation.
SQ SEQUENCE 1150 AA; 129895 MW; 7D6AB93D6BDC97EB CRC64;

Query Match 46.7%; Score 2825.5; DB 1; Length 1150;
Best Local Similarity 48.6%; Pred. No. 2.6e-153;
Matches 574; Conservative 197; Mismatches 364; Indels 47; Gaps 17;

Qy 1 MSPNNQYELIDASSSTSVSDNSVRYPLANDOTTTTLQNNMYKDYLRMSGE-----NP 54
Db 1 MNRNPNNEYIIDAPYCGCFSDDDVRYPLASDPNRAFNQNNYKDYLYQDGYDYGSLNP 60
Qy 55 ELFGNPEPTFSSSTVTQGTIGVQVGLGALGVPFAGQIASPYSFIVGQLWPSSTVSWEM 114
Db 61 NLSINP-----DVLQGTINIVGRILGLFGLVPPFAGQLVPTFTLLNQLMPTNDNAWEAF 115
Qy 115 MKQVEDLIDOKITDTSVTKTALAGLGDLGDLVDTQKSLKNWLNENRNDTRASVVTQYIA 174
Db 116 MAQTEELIDOKISAQVVRNALDDLTGLHDYEEYLAALAEWLRLPNGARA-NLVQRFEN 174
Qy 175 LEQDFVAKISPAIS-CQE---VPLLSVYAQANLHLLLRDASIFGAEWGFTGCEISTP 230
Db 175 LHTAFVTRMPSFGTSGPSGRDVAALLTVYAQANLHLLLRDASIFGAEWGFTGCEISTP 234
Qy 231 YDRQVTRTAQYSDYCVKRWYNTGLDKGTNAASWLKTHQPREMTLLVLDLVALPNNYD 290
Db 235 FNAQQETRIYTHCHVETVYARGLEDVRGTWNTSEWLNHYRFRREMTLMAMDLVALPNNY 294
Qy 291 RTYPIETTAQLTRVYTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLPDILS 350
```

Db 295 ROYFNGANPOLTRREIYTDIVNPPANQICRWGNPNPNTSELENAPRPHPLPERLN 354
Qy 351 EIEFTYTRAGLPLANNTELEYVWGHSHIKYKNTNASSALERNYGTITSNKIKYIDLANKDI 410
Db 355 RLTISSRRYTAPTTNS-FLDYSNGHTLQSHANNPTVYTSYGQITSN-TRLPNTTN-CA 411
Qy 411 FOVRISGLADLANYAQVGPVYASFTLLDKNTGSGVGFTYSKPHHTMVCQVNTID 470
Db 412 RAIDSRARRFGLNYLANLGV--SSLNIPP--TG-----VMSHETNAANTCRQDLITTE 460
Qy 471 EIPPENEPLSRGSHRLSHITSYSPSKNASSP-ARYGMLPVPFAWTHRSADVNTVYSKDI 529
Db 461 ELPLENNP-----NLLSHVTLRFTNTQGGPLATLGPVTVVTRVEDVFTTITADRI 515
Qy 530 TOIPVVKATHTLVGTTVINGPFTGNILKRTSGPLATSVSVKSPISQRYRARIAS 589
Db 516 TQLPVPVKAISEIGGTTVVKPGFTGDLRRTDGGAUGTIRANVAPLTQOYRIRLYAS 575
Qy 590 TTN--LRLPVTISGTRIYSINVKMTWKGGDLTENTEDLATIGTAPTSNYSDSLTVGAD 647
Db 576 TYSFVNLFVNNSAA---GFTLPSTMAQNSLITSEFNTLEVTHTIRFSQSDTTLRLMIP 632
Qy 648 SPASGGEVYVDPKHLIPVNAFEABEDLVAKKAVNELFTSKDALQTSVTDYQVNOAAN 707
Db 633 PSISQGEVYVVDKLEIVNPITREASEDEDAKKAVALFTRTRDGLGVNNTDYQVNOAAN 692
Qy 708 LVECLSDLELYPNEKMLDVAKEARLVQARNLQOTGNRING--ENGWGTSGTGIEVAB 765
Db 693 LVSCLSDEQYGHDKOMLEAARAAKRLSRERLLQDPDFNEINSTEENGWKAASNGVTISE 752
Qy 766 GDVLPKDRSLRITSAREIDETPTVYLVQIDESLLKPVTRKLAGPIGSSQDLKILIR 825
Db 753 GGPFFKGRALQASAR---ENPTYIYQKVDASTLKPTRYKLDGFGVQSSQDLSDILIH 808
Qy 826 HRANOIVKRNPNLFDVLPVNSCGGIDRCSEQQYVDANLALENNENGNGNMSDSSHAFSP 885
Db 809 HKHVHLKVNPNLVNDSYSDGSCGINRCBEOHQVDVQDLDAEDHPKOCCEAAQTHEFSS 868
Qy 886 HIDTGEIDLENWGTIWWFKIPTNGYATLGMLEIVEGPGLSGFTYLERAAQOBSQOQDKM 945
Db 869 YHTGDLASVDQGIWVYLVQRTTDTGATLGNLEIVEGPGLSGSELERQDNKAKNBEV 928
Qy 946 ARKRGASKAYAAQADIRLFPADYQDKLNSGVMSDMLAAQNLVQSIPIVYNDALPRI 1005
Db 929 GKRAETDRIYQDAKAQNLHLPVDYQDQLSPGVGHADILDAQNLIASISDVISDAVLOI 988
Qy 1006 FGMNYSPTFELNRLQOAWNLVDLENAIPNGDFRNGLSDMNATSDVNYQQLSDTSVLVIP 1065
Db 989 PGINYEYMTLSNRLQOASLYTSRNVVQNGDFNSGLDSWNATDTATVQDGNMHEFLVS 1048
Qy 1066 NWNVSQSOQTPQPNRYVLRVLTARKEGVDGYVIRIGANQTELTFTNI CDDDTGVLSA 1125
Db 1049 HMDAQVSOQFRVQPNCKVLRVLTAKKVGNGDGYVTIODGAHRETLTTFNACDYDNGTHV 1108
Qy 1126 DQTSYITKTVETPTGTOVIMDSMTEGVENLKSVELLERE 1167
Db 1109 NDNSTYITKELVPYPTKHEMWEVSETEGTFYIDSTIEPTQE 1150

RESULT 6

C9DA_BACTP STANDARD; PRT: 1169 AA.
AC 006014;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Peatidial crystal protein cry9Da (insecticidal delta-endotoxin
DE CryIXD(a) (Crystalline entomocidal protoxin) (132 kDa crystal
DE protein).
GN Cry9DA OR CryIXD(A).
OS Bacillus thuringiensis (subsp. japonensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

NCBI_TaxID=128936;
OX [1] SEQUENCE FROM N.A.
RN STRAIN=N141;
RA Asano S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D85560; BAA19948.1; -
DR HSP; P07130; 1DLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin_1.
DR Pfam; PF03944; endotoxin C; 1.
DR Pfam; PF03945; endotoxin N; 1.
DR Toxin; Sporulation.
KW Toxin; Sporulation.
SQ SEQUENCE 1169 AA; 132228 MW; 659AB257229DE5E9 CRC64;

Query Match 44.9%; Score 2715; DB 1; Length 1169;
Best Local Similarity 46.9%; Pred. No. 5.3e-147;
Matches 558; Conservative 204; Mismatches 383; Indels 46; Gaps 17;

Qy 1 MSPNNQNYEILDASSSTVSNSVRYPLANDQTTTLQNMNYKYLRMSEGENPELFGNP 60
Db 1 MNRNNQNEYVIDAPHCPCPADDDVVKYPLTDDPNAGLQNMNYKEYLQTYGGDYTDPLNP 60
Qy 61 EFTFS-SSTVGTGIGVQVGLGALGVPFAGQIASFVPIVQGLPSSSTVSVMEMIMKQVE 119
Db 61 NLSVSGKDVQVGINIVGRLSPFPFPSSQWVTVYLLNSLNFDDNSVMDAPMERVE 120
Qy 120 DLIQKITSVRKTALAGLQGLDGLDVYQSKLKNLENRNDTRARSVVVYQYIALELDF 179
Db 121 ELIDQKISEAVKGRALDDLGTGLQVNYLVVEALDEMLRPNRGARA-SLVSRFNILDSLF 179
Qy 180 VAKTSPALSG-----QEVPLLSVYAOAANLHLLILRLDASIECAEWGTFPGHISTPYDRQ 234
Db 180 TQFMPSPG-SGPGSQNYATILLPVYAOAANLHLLILRLDADIIYGARWGLNQTDIQPHSRQ 238
Qy 235 VTRTAQYSDYCVKQVNTGLDKLKGNTAAASLKYHOFREMTLLDLVALFPNYDRTYTP 294
Db 239 QSLQTTYNHCVTAANDGLAELRGTTAESFKYQYRREMTLTAMDVLVALFPYNNLRQYP 298
Qy 295 IETTAQLTRRYVTDPIFVN--RETSGGFCRWLSN-----SDISPEVSASVIRSHPLFDI 348
Db 299 DGTNPQLTRRYVTDPIAFDPLBQPTQLCRSWYINPAFRNHLNFSVLNLSLRPPHLPFR 358
Qy 349 LSEIEP----YTRAGLPLNNTLEYWVGHSHIKYKNTNASSALERNYGTITSNKIKYIDL 405
Db 359 LSNLQILVNYQTGSA-----WRGSRVRYHLYHSSIIQEKSYGLLSDFGVGANTINV 408
Qy 406 ANKDIQVRSGLADLANYAQVGPVYASFTLLDKNTGSGVGFTYSKPHHTMVCQVNTID 465
Db 409 QNNDIYQIISOVSFNPASPVGSSYSVDNFTYI-----SSGQVSGISGTYQOQIPACVLOQ 463
Qy 466 YNTIDEIP---PENEPLSRGSHRLSHITSYSPSKNAS-SPARY-GNLPVFAWTHRSADV 520
Db 466 YNTIDEIP---PENEPLSRGSHRLSHITSYSPSKNAS-SPARY-GNLPVFAWTHRSADV 520

```
Db 464 RNSTDPLSNPEGD-IIRNYSHRLSHITQYRQATQSGSPSTVSANLPTCVWTHRDVLD 522
Qy 521 TMTVYSDKITQIPVKAHALVNSGTTVIKGPFTGNGILKRTSGPLAYTSVSKSPLSQR 580
Db 523 DMTIYANQITQIPVKAHALVNSGTTVIKGPFTGNGILKRTSGPLAYTSVSKSPLSQR 582
Qy 581 YRARIYASTTNLRPLVPTISGTRIYSINNVKNTKNGDDLTFTNFDLATIGTATFPNYSQ 640
Db 583 YRIRFYASTIDFPVTRGGTITINNFRTRTNRQESRYSTRVETPTNFTQSQD 642
Qy 641 SLTVGADSPASGGEVTVDFKELIPVNAPEAEEDLVAKKAV-NGIPTSKEALQTSVD 699
Db 643 IIRTSQISGNGEVEVDRIEIIIPVNPAPAEEDLEAKKAKARQNLFTTRDGLQNVVD 702
Qy 700 YOVQOANLVBCSLDELYPEKRMKLDVAKKALVQARNLQDTGFNRING--ENGWGTG 757
Db 703 YQVQOANLVBCSLDELYPEKRMKLDVAKKALVQARNLQDTGFNRING--ENGWGTG 762
Qy 758 STGIEVABGDVLPKORSRLTSAREIDTETPTLYLYQOIDESELLKPYTRYKLGFTGSSQ 817
Db 763 SNGVTISEGGPPFKGRALQASAR----ENYPTIYQKVDASVLKPYTRYKLGFTGSSQ 818
Qy 818 DLEIKLIRANQIVKQVNDLPLVDLPVNSCGIDRCSEQQYVDANLLENNGE-NGNM 876
Db 819 DLEIDLIHYKHLVKNVNDLPLVDLPVNSCGIDRCSEQQYVDANLLENNGE-NGNM 878
Qy 877 SDSHAFSPHIDTGEIDLNTGICWVFKIPTNGYATLGNLELVEGSLGSETLERRAQ 936
Db 879 AACTHEFSYINTGDLNASVDQGIWVVKVTRTDGVTATLGNLELVEGSLGSETLERRAQ 938
Qy 937 QBOQWQDMARKRGASEKAYAAKQAIIDLFPADYQDKLNGSVGEMSDMLAONLVOSIPY 996
Db 939 DNAKNNAELGRKRAEDRVYLAQKQAIHLFVDYQDKLNGSVGEMSDMLAONLVOSIPY 998
Qy 997 VYNDALPEIYKNTYFTLWELQOANLYDLRNALPNGDERGLSDNATSDVNVQOL 1056
Db 999 VYSDTLQIQINYEIYITELSDRQASLYLTSRNAVQNGDFNSGWNITTDASVQOD 1058
Qy 1057 SDTSVLVIPNWSQVQFTQVQNYRYLVRVTKARKEGVDGVIIRIDGANQETLTFNIC 1116
Db 1059 GNWHFLVSHWDAQVSQLRVNPKCYLVRVTKARKEGVDGVIIRIDGANQETLTFNIC 1118
Qy 1117 DDDTGVLSADQTSYITKTEVPTSTPQVWIDMSGTEGTVNIESVRLVLREE 1167
Db 1119 DYDVNGTYVNDNSYITEVVFVPTKHMVSESESGSFYDSIRPIETQE 1169

RESULT 7
C1KA BACTM STANDARD; PRT; 1215 AA.
ID C1KA BACTM STANDARD; PRT; 1215 AA.
AC Q45715;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cryIka (insecticidal delta-endotoxin
DE CryI(a)) (Crystalline entomocidal protoxin) (137 kDa crystal protein).
GN CRYIKA OR CRYIKA(A) OR CRYI(K).
OS Bacillus thuringiensis (subsp. morrisoni).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1411;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=F190;
RX MEDLINE=96102856; PubMed=8586263;
RA Koo B.T., Park S.-H., Choi S.-K., Shin B.S., Kim J.I., Yu J.H.;
RT "Cloning of a novel crystal protein gene cryIka from Bacillus
RT thuringiensis subsp. morrisoni."
RL FEMS Microbiol. Lett. 134:159-164 (1995).
CC -! FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS. SELECTIVELY TOXIC TO ARTOGEIA RAPAE
CC AND NOT ACTIVE ON PLUTELLA XYLOSTELLA.
CC -! DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
```

```
CC the spore coat.
CC -! MISCELLANEOUS: Toxic segment of the protein is located in the N-terminus.
CC -! SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U28801; AAB00376.1; -.
CC HSPB; P02965; 1C1Y.
CC InterPro: IPR001178; Endotoxin.
CC InterPro: IPR005638; endotoxin_C.
CC InterPro: IPR005639; endotoxin_N.
CC InterPro: IPR008979; Gal_bind_like.
CC Pfam: PF00555; endotoxin_1.
CC Pfam: PF03944; endotoxin_C; 1.
CC Pfam: PF03945; endotoxin_N; 1.
CC Toxin; Sporulation.
CC SEQUENCE 1215 AA; 137378 MW; 7CAP7A3311893D9B CRC64;
Query Match 44.9%; Score 2714; DB 1; Length 1215;
Best Local Similarity 46.3%; Pred. No. 6.4e-147;
Matches 581; Conservative 194; Mismatches 351; Indels 130; Gaps 23;
Qy 1 MSPNNQVEYILDASSTVSQNSVRYPLANDQTTTLQNNYKYDLYMSGENPELPCNP 60
Db 1 MNSNRKNEEIIINALSIPAVNSHQMLSPD-----ARIDSLCVAEG-----NNI 47
Qy 61 ETPISSTVQTGIGVGVGLGVPFAGIAPFYSPVQGLPSSVTVVEMIMKQVED 120
Db 48 DFPVASTVQTGISIAGRILGVLPFAGIAPFYSPVQGLPSSVTVVEMIMKQVED 106
Qy 121 LI-DQKITPSVRKATLAGLGDGLDYQKSLKNLKNLNNRNDTRARVVVTVYIALELDF 179
Db 107 IVRQQTITDSVRDTAARLEGLRGVRSYQQALETYLDNENDARSRIIRERYIALELDI 166
Qy 180 VAKIPSPALSGQEVPLISVYQAANLHLALLRLDASIFGAEWGTPPGISTFYDQVTRTA 239
Db 167 TTAIPFISIRNBEVPLIYVQAANLHLALLRLDASIFGAEWGTPPGISTFYDQVTRTA 226
Qy 240 QYSDYCVKWTGLOKLGKNAASWLKYPHREMTLLVLDLVALFPNYDTRTYPIETTA 299
Db 227 EYSNHCQWNTGLNRLRGTTAETWRYNQFRDLTLGLVLDLVALFPNYDTRTYPIETTA 286
Qy 300 QLTREYVYTD--IVFNRETSGGFCRRWLSNDSISPSSESAVTRSPHLPDLSIEFTTT 357
Db 287 QLTREYVYTD--IVFNRETSGGFCRRWLSNDSISPSSESAVTRSPHLPDLSIEFTTT 339
Qy 358 RAGLPLANTYLEYVWYGHISIKYNTWASSALERNYGTITSNKIKYDYL--ANKDIFQVRS 415
Db 340 RSGV---GTTIMNLWAGHRITFNRIQGGSTSEMYGAIT-NPVSVDIPFNVNDIVYTVS 395
Qy 416 LGADLANYVYQVYVYASFTLLDKNTKGTSGSGVGGFTYSKPHITTMQVCTQNTYIDIPPE 475
Db 396 LAGGLGSLSGIRYGLTRVDVDFMIFRN-HPDIVTGLFVHPGHA--GIATQVKDSDELPE 452
Qy 476 --NEPLSRGYSRHLSHITSYSPSKNASSPARYGNLVPFANTHRASDVNTVYSDKITQIP 533
Db 453 TTEQPNYRAPFSLLSHI-----SMGPTTQVQVFPVYVSWTHOSADRTNTINSRITQIP 504
Qy 534 VYKAHTLVSGTTVIKGPGFTGNGILKRTSGPLAYTSVSKSPLSQRYARIRVASTTNL 593
Db 505 LVKAHTLVSGTTVIKGPGFTGNGILKRTSGPLAYTSVSKSPLSQRYARIRVASTTNL 564
Qy 594 RLFPVITSGTRIYSINNVKNTKNGDDLTFTNFDLATIGTATFPNYSGLTVGADSPASG 653
Db 565 RIYTVAGERIYAGQDFKTMADAGALPTFQSGFSVATINTATFPERSSSLTIGADTFSSGN 624
```



```
Qy 654 EYVYDKFELIPVNAIFAEEDLDVAKKAVNGLFTSKDD-ALQTSVTDYQVNAQANVECL 712
Db 625 EYVYDFELIQVATFRAESDLERAKAVNALFTSTNPGRLKTDVTDYHIDQVSNLVECL 684
Qy 713 SDLEYLPNEKRLMDAVKAEKLVQARNLQDTGNRLNG--ENGWGTSGTGLEVAEGDVL 770
Db 685 SDFCLDKKRELLREVKYAKKLSDERNLQDPTFTSISGQDTRGWTGSIQGGDDIF 744
Qy 771 KDRSLTSAREIDTETETPTLYQOQIDESLLKPYTRYKLGFIQSSDLEIKLIRHRANQ 830
Db 745 KENVRLPGT--VD-ECYPTLYQKIDESQLKSYTRYQRCYVIEDSQDLLEYLYRYNAKH 801
Qy 831 IVKRVNPMLP-----DVLVFNSSCGIDRCSSQOQVVDANLALNNGENGMG-SSSHAFSP 885
Db 802 ETLSPVPGTESPWPSSGVYPSRCGEPNRCAPR--IENPDLDCSCRYGEKCVHSHFSL 859
Qy 886 HIDTGEIDLNENTGIWVVKLPITPTNGVATLNLALVEBGLSGETLERAOQOQWODKM 945
Db 860 DIDVGCIDLNEDLGVWVIFKLTQDGHAKGNLFIEKPLGLGKALSKVRKAERKWDKY 919
Qy 946 ARKRGASEKAYAAKQAIKDLFADYQDQKLSNGVEMSDMLAAQNLVOSIPYVYNALPEI 1005
Db 920 EKLQETKRVYTBKESVDALFVDSQYDKLQANTNIGIIGHGADQVHREIPYLSLPVI 979
Qy 1006 PGMVYTSFTELTNRLQOANLYDLRNLAIKNGDFRNLSDMNTSDVNVQQLSDTSVLVIP 1065
Db 980 PSINAAIFEELEHGIKFRAYSIDYARNVIKNGDFNGLSCNNVKGHDVQVQNHHSVLVS 1039
Qy 1066 MNSVQSQOFTVQPNRYVLTARTKSGVGGVYLIIRDGANOTELTFN----- 1114
Db 1040 EWEAEVSQKRVCPDRGIVLTAYKSGYGGCVTIFHEFQNTDVLKFRNVEHEVYVNN 1099
Qy 1115 --ICDDDTGVLSADQTS----- 1129
Db 1100 TVTCNDYTNQSGSDTACNSYNGVEDYENRYEPNAPSAPVNYTPYRBMVTDQGY 1159
Qy 1130 -----YITKVEFTPTSTQWIDMSHETGVPNIESVELVLBBE 1167
Db 1160 NHCVSDRGYRNHTPLPAGVYVTELEYPETEQVMEIETGRTFVGSVELLMBE 1215

RESULT 8
ID C1BB BACTU STANDARD; PRT; 1229 AA.
AC Q45739;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pesticidal crystal protein cryIb (insecticidal delta-endotoxin
GN CRYIb(b)) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxId=1428;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-21110 / EG5847;
RA Donovan W.P., Tan Y., Jany C.S., Gonzalez J.M. Jr.;
RT "Bacillus thuringiensis cryet4 and cryet5 toxin genes and proteins
RT toxic to lepidopteran insects.";
RL Patent number US5322687, 21-JUN-1994.
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L32020; AAA22344.1; -.
CC HSP; P02965; ICYI.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin C.
CC InterPro; IPR005639; endotoxin N.
CC InterPro; IPR008979; Gal_bind_Like.
CC Pfam; PF00555; endotoxin; 1.
CC Pfam; PF03944; endotoxin; 1.
CC Pfam; PF03945; endotoxin; 1.
CC Toxin; Sporulation.
CC KW Toxin; Sporulation.
CC SQ SEQUENCE 1229 AA; 139769 MW; A4C949DB675C3269 CRC64;

Query Match 42.2%; Score 2552; DB 1; Length 1229;
Best Local Similarity 43.2%; Pred. No. 1.1e-137;
Matches 548; Conservative 205; Mismatches 375; Indels 140; Gaps 22;

Qy 1 MSPNQMEYEILDASSSTVSNDNRYRYPLANDQTTTLQMNMYKDYLRMS EGENPELFGNP 60
Db 1 MTSNRKNEEIIALS IPTVSNPSTQWNLSPD-----ARIEDSLCVAE-----VNNI 47
Qy 61 ETPFSSSTVQGGIGIGVQVIGALGVPPAGQIATSPYSPVIGQIWPSSVSVHEMIMKQVED 120
Db 48 DPPVYASTVQGGINAGRIILGVLPVPPAGQIATSPYSPVIGQIWPSSVSVHEMIMKQVED 106
Qy 121 LIDOKITDVRKTAAGIQLGIDGIDVYOKSKWLENRNDTRARSVVVTVYALIELDPV 180
Db 107 LIRQOVTENTNTAIALLEGIGRGYRYSQQALETWLDNRNDARSIIILERVIALELDIT 166
Qy 181 AKIPSAISQGEVPLLSVYQAANLHLLLRDASIFGAEWGTFCGEISTFYDRQVTRTAQ 240
Db 167 TAILPLRIREEVPLLVYQAANLHLLLRDASIFGSEWGWASSDVNQYQEQIRYTB 226
Qy 241 YSDYCVKWTNGDKLKGNTAASWLKTHQPRREMTLLVLDLVALPNTDTRYPYETTAQ 300
Db 227 YSNHCVQWNTYGLNLRGTNAESWLRYNQPRDLTLGLVLDLVALPNTDTRYPYETTAQ 286
Qy 301 LTRVYVTDPIVFNRETSGPCRRWSLNSDISFSEVESAVIRSPHLFDLSIEFTYTRAG 360
Db 287 LTRSIYTDPIGRTNAPSGFASNTNFWNNAPSFAIEAAIAPPPLLDLPFQLTIYS--AS 344
Qy 361 LPLANNTEYLEYVGHGSIKYK-----NTNASSALERNYGTITSNKIKYDLANKDIFQVRS 416
Db 345 SRNSESQTHMNYVWGHRLNFRPIGGTLNTSTQGLTNTSINPVLQF---TSRDVYRTESN 401
Qy 417 GADLANYAQVGVYPYASFTLLDKNTGSGVGGTYSGEHTTMQVCTQNYNTIDRIPEP- 475
Db 402 AGTNILFTTPVNGVPMFARFNPQ-NIYERGATTSQYQGVGI--QLFDSLELPPET 458
Qy 476 -NEPLSRGYSRLSHITSYSKSNASSPARYGN--LPVPANTHRSADVNTVTVYSDKITQ 531
Db 459 TERPNESYSRLSHIGLI-----IGNTLRAPVYSWTHRSADRTWTIGNRITQ 507
Qy 532 IPVYKAITLVSGTIVIRKPGFTGGNILKRTSSGLPAYSVSVKSPLSQRYRARIYAST 591
Db 508 IPLVKALNLHSGVTVVGPGFTGGDILLRTWTGTGDIRLNLINVLSPSQRYRARIYAST 567
Qy 592 NLRPLVTISGTRIYSINWNTNKGDDLTFTWPLATIGTAFTEPSYSDSLTVGADSPAS 651
Db 568 DLQFPTRINGTIVNIGNFRSRTMNRGDLNLEYSFRTAGSTFPFNFLNAOSTFTLGHQSP-S 626
Qy 652 GGEVYVDKFLIIPVNAIFAEEDLDVAKKAVNGLFTSKDD-ALQTSVTDYQVNAQANVE 710
Db 627 NQEVYIURVEFVPAEYFEAYDYDERAKAVNALFTSTNPRLLKTDVTDYHIDQVSNVA 686
Qy 711 CLSDLEYPNKRLMDAVKAEKRLVQARNLQDTGNRLNG----- 751
Db 687 CLSDFCLDKRELLFEKVYAKRLSDERNLQDPNFTFISGQLSPASIDGOSNPPSINEL 746
```



```

QY 752 -ENGWGTGTCGIRVAGDVLKFORSLRLTSARBIDTETPTLYLQYQIDESLLKPYTRYKLK 810
Db 747 SEHGWSANVTYQEGNDVFKENYVTLPGT---FNECYPNYLYQKIGSELKAYTRYQLR 803
QY 811 GFTGSSQDLKILIRHRANQIVKRNPDNLDPVLVNSCGGIDRCSEQQ---YVDANLA 866
Db 804 GYIENSQDLIYLIRYNAKEAIVNPGT--ESWISAESITIGKTEPNRCAPHYENPD 861
QY 867 LENNGENG-NMSDSDSHAFPHIDTGEIDLNENTGIWVVFKEIPTVNGYATLGNLELVEEGP 925
Db 862 LDSCSRDGEKCAHSHSHSTLIDVGCCTDLHENLGVMLIPKIKTQDGHARLGNLEYLEKP 921
QY 926 LSETTERAQQQQQQQWQDWARTRGSEKAYYAKQAIIDLFPADYQDQKLSNGVMSDML 985
Db 922 LLGEALRVKRTKREKREKLEHLETKRYTYAKESVDALFVDSQYDLQANSNGMIH 981
QY 986 AAOQLVQSIPIVYNDAIPRIKPGANNYSFETLTVRLQOANLYDLRINAIKNGDFRNGLSDW 1045
Db 982 AADKLVHSIREAVLSELVTRGVNADIFELBEGHILTAFLSYDARNAVKNQDFRNGLTCW 1041
QY 1046 NATSDVNVQSLDTSVLVPIPNMSQVSQQFTVQPNRYVLRVTRARKEGVGDGVYIIRDA 1105
Db 1042 NVKGVHDVQCSHRRFDLVPEMKAQVSVQVRVCPGCGYILLRVTRAYKEGEGCVTHIEIB 1101
QY 1106 NQTEITLP-----NICDDDTGV----- 1122
Db 1102 ENTDELNFKNRVBEIYPPDTGTCKYYTENQGTTRTCGNEGSRNEGYDAYEINAKSSLE 1161
QY 1123 -----LSADQTSYTKTVEPTSPSTEOVWIDMSETEGVP 1155
Db 1162 YRPTYBEETVDRRENHCYARGYINYSVPVAGYVTKLEYFPETDVTWIEIGTEGKF 1221
QY 1156 NIESVELMLEE 1167
Db 1222 IVDSVELLAMEE 1233

RESULT 11
CIBE_BACTU
ID_CIBE_BACTU STANDARD; PRT; 1227 AA.
AC 08505;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cryIbE (Insecticidal delta-endotoxin
DE CryIb(e)) (crystalline entomocidal protoxin) (139 kDa crystal protein).
GN CRYIbE OR CRYIb(E) OR 158C2B.
OS Bacillus thuringiensis.
OG Plasmid pMYC2383.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-18872 / PS158C2;
RA Payne J.M., Cummings D.A., Cannon R.J.C., Narva K.R., Stelman S.;
RT "Bacillus thuringiensis genes encoding lepidopteran-active toxins.";
RL Patent number US5723758, 03-MAR-1998.
CC -1- FUNCTION: PROMOTES COLLOIDOSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
CC -1- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -1- SIMILARITY: Belongs to the delta endotoxin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
```

```

CC -----
DR HMBP; AF077326; AAC32850.1; -.
DR HSP30; P07130; 1DLIC.
DR InterPro; IPR001178; Endotoxin.C.
DR InterPro; IPR005638; endotoxin.C.
DR InterPro; IPR005639; endotoxin.N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
KW Toxin; Sporulation; Plasmid.
SQ SEQUENCE 1227 AA; 139084 MW; CBA847BEA0B34CD3 CRC64;

Query Match 40.7%; Score 2462; DB 1; Length 1227;
Best Local Similarity 42.9%; Pred. No. 1.5e-132;
Matches 542; Conservative 196; Mismatches 392; Indels 134; Gaps 20;

QY 1 MSPNNQNEVEILDASSSTVSNSVRYPPLANDQTTLQNNYKYDLRMSRGENPELFCNP 60
Db 1 MTSNRKNNEIINALSIPVNSHSAQNLSSTD-----ARIEDSLCIABG-----NNI 47
QY 61 ETEISSSTVQTGIGIVGQVLGALGVPAGQIASFYSPIVGQLMPSSTVSVMEMIMQVED 120
Db 48 DPFVSASTVQTGINIAGRIILGVLGVPAGQIASFYSPVLGELMPGR-DPWEIIFLEHVEQ 106
QY 121 LIDQKITDSVRKLTALAGLOGLDGLDVYQKSLKNWLENRNDTRARSVVVTOYIALELDFV 180
Db 107 LIRQQVTESTRDALTALARIQGLGNSFRAYQQSLSDWLENRDDARTRSVLYTOYIALELDFL 166
QY 181 AKTPSPAIQKQEPILSVYQAANLHLLLRDASIPGAEWGPTEPGEISTFYDROVTRTAQ 240
Db 167 NAMPLFAIRNQEEVPELVVYQAANLHLLLRDASLPGSEFGLTSQEQRYTERQVEKTR 226
QY 241 YSDYCVKQVNTGLDKLGTNAASWKLKYHQFRREMTLLVLDLVALFPNYDTRTYPIETTAQ 300
Db 227 YSDYCARVNTGLANLGTNAESWLVRYNQFRDLTLGLVLDLVALFSPVDTRVYPMNTSAQ 286
QY 301 LTRVYVTDPIVFNRETSGGFCRWLSNSDISPSEVESAVIRSPHLFDILSIEIFYTTRAG 360
Db 287 LTRVYVTDPIGRTNAPSGFASFWNNAPSPSAIEAAVIRPPELDPPEQLTIFSVLS- 345
QY 361 LPLANTEYLEVYVGHISKYKNTNASSALERNYGTITSNKIKYYDLANKDIFOVSLGADL 420
Db 346 -RWSNTQYIMYVWGHRLSRTIGSLSTSTHGTNTSINPVTLOFTSRDVRTRSPAGIN 404
QY 421 ANYTAQVGVYASFTLLDKNTGSGVGGFTYKSPHTTMQVCTQNYNTIDIPPE--NBP 478
Db 405 ILLTTPVNGVFWARFVW--RNPLNSLRGSLTYTGYT--GVGTQLPDSSETLPETTERP 460
QY 479 LSRGYSHRLSHITSYSPSKNASSPARYGN--LPVFANTHRSADVTTVYSDDKITQIPVV 535
Db 461 NYESYSHRLSNIRLIS-----GNTLRAPVYSWTHRSADRTNTTSSDSITQIPLV 509
QY 536 KAHTLVSGTGVTKRGPTGGMILKRTSSGPLAYTSVSVKSPLSQRYRARIYASTTMLRL 595
Db 510 KSPNLNSGTYSVSGSGPFTGGDIIRTNVNGSVLMSGLNFNNTSQRYRVRVRYAASQTMVL 569
QY 596 FVTISGTRIYSINVNKTMKGGDLTNTFTDLATIGTFTFTSNYSDSLTGVGADSPASGGEV 655
Db 570 RVTVSGSTTFDQGFPTSMANESLTSQSFPAFPFVIGISAG-SQTAGISINNAGRTTP 628
QY 656 YVDKFEIIPVNAITFEAEEDLDVAKAVNGLFTSKDD-ALQTSVTVDYQVNAQANLVECLSD 714
Db 629 HPDKIEFIPITATFEAEYDLERAQEAVALFTNTNPRRLKATGVTVDYHIDEVSLVACLSD 688
QY 715 ELYPNKRMLDVAKEAKRLVQARNLLQDGFVRIN-----GNG 754
Db 689 EFCIDSEKRELEKVKVAKRLSDERNLLQDPNFTSINKQDPFITNEQSNFTSIHQSEHG 748
QY 755 WTGSGTGLVABGVDLFKDRSLRLTSAREIDTETVPTLYLQYQIDESLLKPYTRYKLKFIG 814
Db 749 MWGSENIITQEGNDVFKENYVILPGT---FNECYPTLYQKIGAEKAYTRYQLSGVIB 805
```

```
Qy 815 SSQDLKIKLIRHRANQIVKVP--DNLLPDVL--PVNSCGGIDRCSEQQVYDANIALENN 870
Db 806 DSQDLKILYLRNNAKHETLDVGEFVSVMPLSVESPIGRCEPNRCA--PHFENPDLDCS 863
Qy 871 GENG-NMSSDSHAFSFDITCEIDNENGTGWWVFKIPTNGYATLGNLELVESGPILSGB 929
Db 864 CRDEKCAHSHHPSLDDVGCIDHENLGVVVFVKIKTQSHARLGNLEFIEKPLLGE 923
Qy 930 TLRAQOQOQWQDMARKGASKAYAAQOADRPLADYQDQKLSGVESDMKLAQN 989
Db 924 ALSVRAEKWRDKREKLEKTRVYTAKEAVDALFVDSQYDRLOADTNGIMHAADK 983
Qy 990 LVQSIPIVYNDALEIPICMNYTSFTELNRLLQOANLYDLRANLPGNGFRGLSDMNAIS 1049
Db 984 LVHRIRAYLSVLPVGNABIEFELEGRITLISLYDARNVKGNGDFNGLACMVKG 1043
Qy 1050 DVNYQQLSDTSVLVPIPNNSQVQFTQPNRYVLRVLTARKEGVGQYVIIRDCANQTE 1109
Db 1044 HVDVQOQSHRSVLVPIPEWAEVSQAVRCPGRGILRLVTAKEGVGEGCVTIHEIENYTD 1103
Qy 1110 TLTRNICDD-----DTGVLS-----ADQTS----- 1129
Db 1104 ELKFENCEEEVPTDTGTCNDYTAHQTAACNSRNAGYDAYEVDVDTASVNYKPTVYEE 1163
Qy 1130 -----YTKVTEPTSTEQVWIDMSFEGVNIYESVELV 1163
Db 1164 TYTVRRDNHCYDRGYVNPVPAGYMTKEVFPFDTKVMIEIGETBGRFVDSVELL 1223
Qy 1164 LEEB 1167
Db 1224 LMBE 1227

RESULT 12
ID C7AA BACTU STANDARD; PRT; 1138 AA.
AC Q03749;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pesticidal crystal protein cry7Aa (Insecticidal delta-endotoxin
DE CryVIIA(a)) (Crystalline entomocidal protoxin) (129 kDa crystal
DE protein)
GN Cry7AA OR CryVIIA(A) OR CryIIIC.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=92384571; PubMed=1514800;
RA Lambert B., Hofte H., Annys K., Jansens S., Soetaert P., Peferoen M.;
RT "Novel Bacillus thuringiensis insecticidal crystal protein with a
RT silent activity against coleopteran larvae."
RL Appl. Environ. Microbiol. 58:2536-2542 (1992).
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF COLEOPTERA. THIS PROTEIN IS NOT TOXIC IN ITS
CC NATURAL FORM. IT IS HIGHLY TOXIC TO COLORADO POTATO BEETLE LARVAE
CC AFTER AN IN VITRO SOLUBILIZATION AND TRYPSIN ACTIVATION STEP.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; M64478; AAA22351.1; -.
DR EMBL; A07236; CA000646.1; -.
DR PIR; A48944; A48944.
DR HSSP; P07130; 1DLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
KW Toxin; Sporulation.
SQ SEQUENCE 1138 AA; 129391 MW; 69D8676D4F6A1FAC CRC64;

Query Match 40.7%; Score 2458.5; DB 1; Length 1138;
Best Local Similarity 43.5%; Pred. No. 2.2e-132;
Matches 514; Conservative 208; Mismatches 400; Indels 59; Gaps 17;

Qy 1 MSPNQNEVEYILDASSSTVSVDNSVVRVPLANDOTTLLQNMNVKDYLRMSGEENBELPGNP 60
Db 1 NMLNLDGYE-----DSNRTLANSNLNPTQKALSPSLKMNQYDPLSITEREQPEALASG 55
Qy 61 ETFISSSTVQGTGIGVGOVLGALGVPPAGQIASFYSPFVQGLWPSSTVSVMMIMKQVED 120
Db 56 NT-----AINTVSVGTATLSALGVPGASPTNFKIAGLLMPENG-KINDEFMTEVEA 109
Qy 121 LIDQKIDSVRTKALAGLQGLDGVYQKSLKNWLENRNDTRASVVVTVIALELPV 180
Db 110 LIDQKIBEVNRKAIABLDGLGSALDKYQKALADWLQKDDPEALISVATEPRIIDSLFE 169
Qy 181 AKIPSPALSGORVPLLSVYQAANLHLLALLDASIFPGAENGFTPCGEISTFFYDQVTRTAQ 240
Db 170 FSPSPFKVTGYEIPLLTVYAANLHLLALLDSTLYGDKWGTQNNIEENYRQKRISE 229
Qy 241 YSDYCVKYNNTGLDKLKGNTAASMLKHQFRRMTLLVLDLVALFPNTYTRTYPIETTAQ 300
Db 230 YSDHCTKWNYSGLSLNGSTVEQMINVNRFRREMILMALDLVAVFPDPRYSMETSTQ 289
Qy 301 LTRVYVTDPIVNRNRTSGCGPCRRWSLNSDI--SPSEVSASVIRSHPLFDILSEIEFYTR 358
Db 290 LTRVYVTDVPSLSIS-----NPDIGSPSQMENTAIRTPHLVDLYDELYIYTSK 338
Qy 359 AGLPLNNTVE-YLEYVWGHSIKYQNTNASSALERNYGTITSNKIK--YVDLANKDIFOVRS 415
Db 339 YKAFSHEIQDPLFTWSAHKVSFKSEQNLVTTGYIGTSGYISSGAKVSHGNDIYRTLA 398
Qy 416 LGADLANIYAQVGVYPYASFTLLDKNTGSGVGGFTYSKPHHTMQVTCQNTYNTIDIPPE 475
Db 399 APSVVVYPTQNYGVEQVEFY-----GVKGVHYRGDNKYDL---TYDSIDQLPPD 446
Qy 476 NEPLSRGSHRLSHITSYSPSKNASSPARYGN--LPVFAWTHRSADVNTVTSYDQKITQP 533
Db 447 GEPIHEKTYHRLCHATAI-----FKSTPDYDNATIPFSWTHRSABYNYRIYPNKITKIP 501
Qy 534 VVKRAHTLVSGTTVLKPGFTGNNILKRTSSGSLAVTSVSKPSLSQVRARIRYASTTNL 593
Db 502 AVKMYKLDPSVTVAGPGFTGDLVWRGSTGVIGIKATVNSPLSKQKRVRYRYATNVSG 561
Qy 594 RLFTVTSIG-----TRIYSINVKNTMKNKDDTLFTNTFDLATITGTAFTSPNSYDLSVGADS 648
Db 562 QFNVVNDKITLQTKFQ--NTVETIGEGKDLTYSGFGVIEYSTTIQFPDPEHKPITLHLSLD 619
Qy 649 PASGGEVTVDFELIPVFNATFEAEEDLVAKAVNGLFTSKKDALQTSVTDYQVQNAHL 708
Db 620 LSNSSSFVDSIEPTFPVDVNTAEKEKFAKAVNTLFTFEGNALQKQDVTYKVDQVSL 679
Qy 709 VECLSDLEYLPNEKMLNDADVAKEAKELVQARNLQDTGENRING--ENGWTSVTGLVARS 766
Db 680 VDCISGDLVPEKREKQLNLVYAKLSYSRNLLODPTFDSINSSSENGWYSGNIGVING 739
Qy 767 DVLFRDRSLRLTSAREIDTETPTTYLYQOIQIDESLLKPYTRYKLGFGISSQDLKLIH 826
Db 740 DPFVKGNYLISGTN--DTQ-YPTLYQKIDESKLKEYTRYKLGFGISSQDLVAVIRY 796
```


Search completed: June 21, 2004, 13:37:53
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:34:30 ; Search time 56 Seconds
(without alignments)
6575.174 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYILDASSSTSV.....MSTEGVFNISSVELVLEE 1167

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mbc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5303.5	97.7	1144	2	Q8KZL7
2	2750.5	45.5	1144	2	Q45745
3	2545.5	42.1	1228	2	Q93T75
4	2539.5	42.0	1228	2	Q93NM5
5	2433	40.3	1231	2	Q8KNV2
6	2056.5	34.0	1155	2	Q9F296
7	2050	33.9	1180	2	Q9S5V8
8	2048	33.9	1176	2	Q7WZT9
9	2045	33.8	1176	2	Q45736
10	2034	33.7	1176	2	Q9RC30
11	2032.5	33.6	1155	2	Q93T21
12	2016.5	33.4	1171	2	Q06894
13	2003	33.1	1169	2	Q8GHE8
14	1997	33.0	1189	2	Q9L877
15	1989.5	32.9	1174	2	Q45749
16	1953.5	32.3	1160	2	Q93TF9

17	1948.5	32.2	1128	2	Q9FDC0	Q9fcd0 bacillus th
18	1939.5	32.1	1177	2	Q8GLY5	Q8gly5 bacillus th
19	1937	32.0	1176	2	Q9S514	Q9s514 bacillus th
20	1933.5	32.0	1177	2	Q03743	Q03743 bacillus th
21	1933	32.0	1178	2	Q45768	Q45768 bacillus th
22	1931	31.9	1178	2	Q9R826	Q9r826 bacillus th
23	1930.5	31.9	1177	2	Q45735	Q45735 bacillus th
24	1925.5	31.9	1118	2	Q9AM83	Q9am83 bacillus th
25	1897.5	31.4	1118	2	Q9AM82	Q9am82 bacillus th
26	1836.5	30.4	1280	2	Q8VUK9	Q8vuk9 bacillus th
27	1780.5	29.5	1236	2	Q93T3	Q93t3 bacillus th
28	1746.5	28.9	1118	2	Q9AM81	Q9am81 bacillus th
29	1730.5	28.6	719	2	Q9F0F8	Q9f0f8 bacillus th
30	1727.5	28.6	719	2	Q8S796	Q8s796 bacillus th
31	1717.5	28.4	719	2	Q93NJ5	Q93nj5 bacillus th
32	1702.5	28.2	719	2	Q8KY61	Q8ky61 bacillus th
33	1702	28.2	1254	2	Q8VUL0	Q8vul0 bacillus th
34	1577.5	26.1	1270	2	Q8VUL1	Q8vul1 bacillus th
35	1378.5	22.8	638	2	Q87654	Q87654 bacillus th
36	1278.5	21.2	1155	2	Q9AM80	Q9am80 bacillus th
37	1254	20.7	645	2	Q9S603	Q9s603 bacillus th
38	1245	20.6	652	2	Q9S6N9	Q9s6n9 bacillus th
39	1171	19.4	723	2	Q9S4B5	Q9s4b5 bacillus th
40	1094	18.1	489	2	Q8KNV1	Q8knv1 bacillus th
41	1069.5	17.7	533	2	Q7X3F6	Q7x3f6 bacillus th
42	1061	17.6	381	2	Q45740	Q45740 bacillus th
43	1052.5	17.4	558	2	Q8VW63	Q8vw63 bacillus th
44	1042.5	17.2	526	2	Q32308	Q32308 bacillus th
45	1041	17.2	620	2	Q45720	Q45720 bacillus th

ALIGNMENTS

RESULT 1
Q8KZL7 Q8KZL7 PRELIMINARY; PRT; 1144 AA.
AC Q8KZL7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cry8 protein.
GN Cry8.
OS Bacillus thuringiensis (subsp. galleriae).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29338;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SDS-502;
RA Asano S., Yamamoto T.;
RT "a novel cry8 gene highly toxic to Anomala cuprea.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089299; BAC07226.1; -
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005639; endotoxin C.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
SQ SEQUENCE 1144 AA; 128060 MW; 98F93070C49014AB CRC64;

Query March 97.7%; Score 5903.5; DB 2; Length 1144;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1143; Conservative 0; Mismatches 1; Indels 23; Gaps 1;
QY 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTLQNNMYKYDYLRMSEGENPELFCNP 60
DB 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTLQNNMYKYDYLRMSEGENPELFCNP 60
QY 61 EFTFISSTVGTGIGVQVILGALGVFPFAGQIASPVFVQGLMPSSTSVSWEMIMKQVED 120

```
Db 1118 TEQVWIDMSFEGVFNIESVELVLEEE 1144
RESULT 2
Q45745 PRELIMINARY; PRT; 1144 AA.
ID Q45745
AC Q45745;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DB Delta-endotoxin (Fragment).
GN CRYX GENE.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94085596; PubMed=8262221;
RA Shevelov A.B., Svarinsky M.A., Karasin A.I., Kogan Y.N.,
RA Chestukhina G.G., Stepanov V.M.;
RT "Primary structure of the cryX*-the novel Delta-endotoxin-related
RT gene from Bacillus thuringiensis esp. galleriae.";
RL FEBS Lett. 336:79-82(1993).
DR EMBL; X75019; CAA52927.1; -.
DR HSSP; P07130; IDIC.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
FT NON TER
SQ SEQUENCE 1144 AA; 129399 MW; 7D28594A19C4B065 CRC64;
Query Match 45.5%; Score 2750.5; DB 2; Length 1144;
Best Local Similarity 48.7%; Pred. NO. 2.8e-148;
Matches 562; Conservative 197; Mismatches 341; Indels 53; Gaps 18;
QY 41 NYKDYLRSEGENPELPCGNPETFISSSTVQVIGVIGVGVGLGALGVPPAGQIASPYSFIVG 100
DB 19 SYKDYLRSEGENPELPCGNPETFISSSTVQVIGVIGVGVGLGALGVPPAGQIASPYSFIVG 76
QY 101 QLPSSSTVSVVEMIMKQVEDLIDQKIDTSVRKATLAGLQGLDGLDVYOKSLKMLNENR 160
DB 77 FLNFSNDQAVWEAFIEQKEBLEEQRIISQVVRTALDGLTGIONTYNQYLALKEWEERPN 136
QY 161 DTRARSVVTVQVIALELDFVAKIPSAISGQ-----EVPILLSVYAAQANLHLLLDASI 215
DB 137 GVRA-NLVLPQFEILHALFVSSMPSFG-SGPGSQRFQQLVYVYAAQANLHLLLDASEK 194
QY 216 FGAEGWPTPGRISTPYDRQV-TRTAQYSDYCVKVTNTGLDKGTNAASWLKYHOFRRM 274
DB 195 YGARGLRRESQGNLYFNEIQRTRDYTHNCVNAINGLAGURGTSAESWLKYHOFRRRA 254
QY 275 TLLVLVDLVALFPNTYTRTYPIETTAQLTRVYTDPIVFNRETSGG-----CRRWSLNSDI 330
DB 255 TLMAMDLLALFPYNTTRYPPIAVNPQLTRVYTDPLGVPSSESSLFPELCLRWQTSAM 314
QY 331 SPSEVESAVIAPSPHLPDILSIEPIETTRAGLPLNNTYLYEYVGHSHI--KYKTNASSAL 388
DB 315 TFSNLENAIISPHLPDITNNLMYTGSPFVHLTN-QLIEGHVSHVSSLLASGPTTVL 373
QY 389 ERNVTGTTISNKIKYVDLANKDIFQVRS--LGADLANVYAOVGVVGVVPSFTLLDKNTGSG 445
DB 374 RNNYGSSTTS-IVNTFSPNDRDYQINTSRHTGLGQN--APLFGITRAQF----- 420
QY 446 SVGGFTVSKPHTTMQVCTQNTYNTIDEIP--PENEPISRGYSHRLSHITSYSF-----SK 497
DB 421 -YPGGTYSVTQNALTCQNYNSIDELPSLDNPEPISRSYSHRLSHITSYLRHVTIDGI 479
```

```
Db 61 ETPFISSSTVQVIGVIGVGLGALGVPPAGQIASPYSFIVGQLMPSSTVSVEMIMKQVED 120
QY 121 LIDOKIDTSVRKATLAGLQGLDGLDVYOKSLKMLNENRNDTRARSVVTVQVIALELDFV 180
DB 121 LIDOKIDTSVRKATLAGLQGLDGLDVYOKSLKMLNENRNDTRARSVVTVQVIALELDFV 180
QY 181 AKIPSPAISSQGEVPLLSVYAAQANLHLLLDASIFGAENGFTPGIEISTFYDRQVTRPAQ 240
DB 181 AKIPSPAISSQGEVPLLSVYAAQANLHLLLDASIFGAENGFTPGIEISTFYDRQVTRPAQ 240
QY 241 YSDYCVKVTNTGLDKLKTNAASWLKYHOFRRMNTLLVLVDLVALFPNTYTRTYPIETTAQ 300
DB 241 YSDYCVKVTNTGLDKLKTNAASWLKYHOFRRMNTLLVLVDLVALFPNTYTRTYPIETTAQ 300
QY 301 LTRVYTDPIVFNRETSGPCRRWSLNSDISPSEVESAVIRSPHLPDILSIEPIETTRAG 360
DB 301 LTRVYTDPIVFNRETSGPCRRWSLNSDISPSEVESAVIRSPHLPDILSIEPIETTRAG 360
QY 361 LPLANNTEYLEYVWGHSHIKYKTNASSALERNYGTITTSNKIKYDILANKDIFQVRSIGADL 420
DB 361 LPLANNTEYLEYVWGHSHIKYKTNASSALERNYGTITTSNKIKYDILANKDIFQVRSIGADL 420
QY 421 ANYAQQVGVYASFTLLDKNTGSGVCGFTYSKPHHTMQVCTQNTYNTIDEIPPEPELS 480
DB 421 ANYAQQVGVYASFTLLDKNTGSGVCGFTYSKPHHTMQVCTQNTYNTIDEIPPEPELS 480
QY 481 RGYSHRLSHITSYSPSKNASSPARVGNLPVPAWTHRSADVNTVYSDKIQIIPVVKANTL 540
DB 481 RGYSHRLSHITSYSPSKNASSPARVGNLPVPAWTHRSADVNTVYSDKIQIIPVVKANTL 540
QY 541 VSGTIVIKGPGFTGNNILKRTSSGGLAYTSVSVPKSLQSVRVARIRYASTTNLRLFVTIS 600
DB 541 VSGTIVIKGPGFTGNNILKRTSSGGLAYTSVSVPKSLQSVRVARIRYASTTNLRLFVTIS 600
QY 601 GTRIYSINVKTNMGDDLTNTFDLATIGTAPTFSNYSDSLTVGADSPASGEVYVDKP 660
DB 601 GTRIYSINVKTNMGDDLTNTFDLATIGTAPTFSNYSDSLTVGADSPASGEVYVDKP 660
QY 661 ELIPVNAATPEAEEDLDVAKVAVNGLFTSKDQALQTSVDYQVQVNAANLVECLSDLYPNE 720
DB 661 ELIPVNAATPEAEEDLDVAKVAVNGLFTSKDQALQTSVDYQVQVNAANLVECLSDLYPNE 720
QY 721 KRLMDVAKEAKRLVQARNLLQDTGFNRINGENGTGSGIEVAAGDVLFXDRSLRLTSA 780
DB 698 KRLMDVAKEAKRLVQARNLLQDTGFNRINGENGTGSGIEVAAGDVLFXDRSLRLTSA 757
QY 781 REIDTETTPYLYQOIDSLLKPYTRYKLGPIGSSQDLKILIRHRANQIVKNVPDNL 840
DB 758 REIDTETTPYLYQOIDSLLKPYTRYKLGPIGSSQDLKILIRHRANQIVKNVPDNL 817
QY 841 PDVLVNSCGGIDRCSQOYVDANLALNNGNGNMSDSHAPSFIOTGEIDLNEMTG 900
DB 818 PDVLVNSCGGIDRCSQOYVDANLALNNGNGNMSDSHAPSFIOTGEIDLNEMTG 877
QY 901 WVYFKIPTNGYATLGNLELVEBGLSGETLERAQQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 960
DB 878 WVYFKIPTNGYATLGNLELVEBGLSGETLERAQQOQOQOQOQOQOQOQOQOQOQOQOQOQ 937
QY 961 QAIDRLPADYQDOKLNSGVHSDMLAAQNLVQSIPIVYNDALPEIPGNMYSFTLITNRL 1020
DB 938 QAIDRLPADYQDOKLNSGVHSDMLAAQNLVQSIPIVYNDALPEIPGNMYSFTLITNRL 997
QY 1021 QOAMNLYDLRINAI PNGDFRNLGSDNNATSDVNVQQLSDTSVLVIPNNWSQSQFTVQPN 1080
DB 998 QOAMNLYDLRINAI PNGDFRNLGSDNNATSDVNVQQLSDTSVLVIPNNWSQSQFTVQPN 1057
QY 1081 YRYVLVTRARKEGVGDGVYIIRDGANQTETLTFNCDDDTGVLSADQTSYITKTVEFTPS 1140
DB 1058 YRYVLVTRARKEGVGDGVYIIRDGANQTETLTFNCDDDTGVLSADQTSYITKTVEFTPS 1117
QY 1141 TEQVWIDMSFEGVFNIESVELVLEEE 1167
```



```
Db      856  DLDCSRDGEKCAHSHHPSLDIDVGCTDLHENLGVVWVFKIKTQSGHARLGNLEFIEEK  915
Qy      925  PLSGETLERAQQOQOQWDEWARKGASEKAYAAKADLDELADYQODOKLNSGVENSMD  984
Db      916  PLIGELSRVGRGKWRDREKLEKTRVYTAKEAVDALFVDSQYDRLOADTNIGMI  975
Qy      985  LAAQNLVQSIPYVYNDALPEIPGMNYSFTBLTNRLQOAWNLIDLNAIPNGDFRNLSD  1044
Db      976  HAADKLVRHIREAYLSELFPVPGYNABIFEELEGHIIAISLYDARNVVKNGDFNGLTC  1035
Qy      1045  WNATSDVNVQOLSQTSVLVVIWNNSQVSQQTVPQNYRVLRVTAKEGUGGVYIRDG  1104
Db      1036  WNVKGHDVQOQSHRSDLVIFEWAEVNSQAVRVCPGCGYILRVAYKEGYGEGCVTHBI  1095
Qy      1105  ANQTEITLF-----  1113
Db      1096  ENNTDELKFKRBEHEVPTDGTGTCNDYTAHQGTAGCADACNSRNAGYEDAYEVDVTASV  1155
Qy      1114  -----NICDDDTGVLSD--QTSYIKTVETFTSTQVWIDMSRETV  1154
Db      1156  NYKTYEETVTVRRDNHCEYDRGVNYPVPVAGYVTKLEYFPETDVTWIEIGETEGK  1215
Qy      1155  FNIESVELVLEER  1167
Db      1216  FIVDSVELLMEE  1228

RESULT 4
Q93NM5  PRELIMINARY; PRT; 1228 AA.
ID Q93NM5
AC Q93NM5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DI 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CryiBa.
GN CryiBa.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A., Huang D.;
RA Zhang J., Song F.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF368257; AAK63251.1;
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin C; 1.
DR Pfam; PF03945; endotoxin N; 1.
SQ SEQUENCE 1228 AA; 139566 MW; E86D9842341FB439 CRC64;

Query Match 42.0%; Score 2539.5; DB 2; Length 1228;
Best Local Similarity 43.8%; Pred. No. 3.5e-136;
Matches 557; Conservative 183; Mismatches 382; Indels 151; Gaps 21;

Qy      1 MSPNNQBYBILDASSSTVSNSRVRYPLANDQTTTLQNMKYDKYLRMSBGENPELFGNP  60
Db      1 MTSNRKNEIINAVNSHA-----QMDLLPDARIEDSLCIAEG-----NNI  42
Qy      61  ETPFSSSTVQTGIGVQVGLNIGVPPAGIATSFYSFVIGOLWPSSTVSVMEMIMQVED  120
Db      43  DPFVSASTVQTGINIAGRILGVLPVPPAGIATSFYSFVIGLWPRGR--DQWEIIFLEHVEQ  101
Qy      121  LIDOKITDSVRTALAGLQGLDGLDVYOKSLKNWLENRNDTRASVVVTVQVIALELDFV  180
Db      102  LINOITENARTALRQLGLDSTRAYQQSLEDMLNDRDARISVLVTVQVIALELDFL  161
Qy      181  AKIPSAISQGVPLLSVYAQANLHLLLRDASIFGAEWGFTPGCEISTFYDQVTRTAQ  240
```

```
Db      162  NAMFLFAIRNOVEPLLVYIAQAANLHLLLRDASIFGSEFGLTSQBIQRIYERQVERTRD  221
Qy      241  YSDYCVKYNVTGLDKLGTNAASLKYHOFREMTLLVLDLVALPNTDTRTYPIETTAQ  300
Db      222  YSDYCVKYNVTGLNLSRGTNAASVVRNQFRDLTLGVLDLVALPSPDYTRYPIINTSAQ  281
Qy      301  LTRVYVTDPIVFNRETSRGFCRRWLSNSDISFSSEVESAVIRSPHLPLILSEIEFTTRAG  360
Db      282  LTRVYVTDALIGATGVNWSM--NMYNNAPSFSAIEAAAIIRPHLLDLFLEQLTIFS--AS  337
Qy      361  LPLANTEVLEYWVCHESIKYK-----NTNASSALLERNYGTITSNKIKYYDLANKDIFQVR  414
Db      338  SRWNTKHTMYWRKGTIQSRPIGGGLNTSGATNTSINPVTLR-----PASRDVYRTE  391
Qy      415  SLGADL--ANYAAQVYGVYPYASFLLDKNTSGSGVGETYSGKPHHTMQVQNTYNTIDBI  472
Db      392  SYAGVLLWGIYLEPIHGVPVTFNPNQISDR--GTANYSOPYESPGQLKDSSET--EL  448
Qy      473  PPE--NEPLSRGYSHRLSHITSYSPKSNASSPARYGNLPVFAWTHRSADVNTVYSDKIT  530
Db      449  PPETTERPNYESYSHRLSHIGIILQSR-----VNPVYSWTHRSADRNTTIGPNRIT  500
Qy      531  QIPVVKATHLVSGTTVIKGPGTGTGNTILKRTSSGPLAYTSVSKSPLSORVEARIRYAST  590
Db      501  QIPMKASBELPQGTIVVRGPGTGGDILRRNTTGGFGPIRVTVNGPLTQRIYRIGFYAST  560
Qy      591  TNLRLFVTTISGTIYISNVNKNMKGDDLTENTFDLATIGTAFTFSNTSDSLTVGADSPA  650
Db      561  VDPDFVSRGGTTVNNFRFLATMNSGDELKGNFVRRAFTPTPTTQIOTIIRTSTQGLS  620
Qy      651  SGGEVYVDKFLIPVNAITFEABEDLDVAKAVNGLFTSKKD--ALQTSVTDYVNOQANLV  709
Db      621  GNGEYVVDKIBIIIVTATFRAEYDLERAQAVNALFTNTNPRRLKTDVTDVHIDQVSNLV  680
Qy      710  ECLSDLEYPREKMLDWAKEKRLVQARNLLOUTGPNRIN-----  750
Db      681  ACLSDLEFCLDEKRELLEKRYAKRLSDERNLQPNFTSINKQDPPISTNEOSNFTSIHE  740
Qy      751  -GENGTGTCGIEVARGDVLFKDRSLRLTSAREIDTETYPYLYQO1DESLLKPYTRYKL  809
Db      741  QSEHGWSGSENIITQEGNDVFNKENVYTLPGT---PNECYPTVLYQKIGESSELKAVTRYQL  797
Qy      810  KGTIGSSQDLBIKLIIRHANOIVKNVP--DNLLPDVL--FVNSCGGIDRCSEQQVVDANL  865
Db      798  RGYIEDSQDLBIYLIYIRYNAKHETLDVPGTESLWPLSVESPIGRCEGPNRCA--PHFEMNP  855
Qy      866  ALEKNGENG--NMSSDSHAPSFIHDTGEIDLANENTGIWVVFKIPTTNGVATLGNLHVRG  924
Db      856  DLDCSRDGEKCAHSHHPSLDIDVGCTDLHENLGVVWVFKIKTQSGHARLGNLEFIEEK  915
Qy      925  PLSGETLERAQQOQOQWDEWARKGASEKAYAAKADLDELADYQODOKLNSGVENSMD  984
Db      916  PLIGELSRVGRGKWRDREKLEKTRVYTAKEAVDALFVDSQYDRLOADTNIGMI  975
Qy      985  LAAQNLVQSIPYVYNDALPEIPGMNYSFTBLTNRLQOAWNLIDLNAIPNGDFRNLSD  1044
Db      976  HAADKLVRHIREAYLSELFPVPGYNABIFEELEGHIIAISLYDARNVVKNGDFNGLTC  1035
Qy      1045  WNATSDVNVQOLSQTSVLVVIWNNSQVSQQTVPQNYRVLRVTAKEGUGGVYIRDG  1104
Db      1036  WNVKGHDVQOQSHRSDLVIFEWAEVNSQAVRVCPGCGYILRVAYKEGYGEGCVTHBI  1095
Qy      1105  ANQTEITLF-----  1113
Db      1096  ENNTDELKFKRBEHEVPTDGTGTCNDYTAHQGTAGCADACNSRNAGYEDAYEVDVTASV  1155
Qy      1114  -----NICDDDTGVLSD--QTSYIKTVETFTSTQVWIDMSRETV  1154
Db      1156  NYKTYEETVTVRRDNHCEYDRGVNYPVPVAGYVTKLEYFPETDVTWIEIGETEGK  1215
Qy      1155  FNIESVELVLEER  1167
Db      1216  FIVDSVELLMEE  1228
```


DR	GO: 0015070; F: toxin activity; IEA.
DR	GO: 0006952; P: defense response; IEA.
DR	InterPro: IPR001178; Endotoxin.
DR	InterPro: IPR005638; endotoxin C.
DR	InterPro: IPR005639; endotoxin N.
DR	InterPro: IPR008979; Gal_bind_like.
DR	Pfam: PF00555; endotoxin; 1.
DR	Pfam: PF03944; endotoxin C; 1.
DR	Pfam: PF03945; endotoxin N; 1.
DR	SEQUENCE 1155 AA; 130557 MW; 5D69E3E2F527749D CRC64;
QY	Query Match 34.0%; Score 2056.5; DB 2; Length 1155;
DB	Best Local Similarity 39.2%; Pred. No. 1.2e-108;
DB	Matches 477; Conservative 184; Mismatches 389; Indels 167; Gaps 29;
QY	53 NPGL-----FCNPET--PISSSTVQTG---IGI---VGQVLGALGVPPAGQIASFYSP 97
DB	4 NPNINECIPYNCNLNPNPEVVLGGERIEFTGYPIDISLSLTQFLISFVPGAGVGLGLVDI 63
QY	98 IVGQLMPSSTVSVMIMKQVEDLIDQIKTDSVRKTLAQLGGLDGLDVYVQKSLKMWLE 157
DB	64 IWGIFGPSQ---WDAFLVQIEQLINQRIEBSFARNQAIISRLLEGLSNLVQIVAESFREWEA 119
QY	158 NRNDTRARSVVVQYIALELDVFAKIIISPATSGQEVPLLSVYQAANLHLILLDASIFG 217
DB	120 DPTNPALRBEEMRIQFNDMNSALTTAIFLFAVQNYQVPLLSVYQAANLHLISLVRDSVSPG 179
QY	218 AEWGTFGEISTEYDVRQVTRTAQYSDVCVKWNTGLDKLKGTAASMLKYHQPFRREMTLL 277
DB	180 QRWGQDAATINSRYNDLTRLIGNTDIAVRWNTGLERWGPDSRDRIYMQPARELTLT 239
QY	278 VLQIALFPNYDRTYPIETTAQLTREYVTDPIVFNRETSGGFCRRSLNSDIFSSEVES 337
DB	240 VLDIVSLFPNYSRTYPIRTVSQLTRREIYNPVLNPD--GSP--RGSAGQ-----IEG 289
QY	338 AVTRSPHLFDILSIEPIFTTRAGLPLANNEYLETWVGHSHK-----YKNTNASSALEKNY 392
DB	290 S-IRSPHLMDILNSITITYT-----DAHRGEY--YWSGHQINASVPVSGSPGPEFTFPL---Y 338
QY	393 GTTISNKIKYVDLANKDIPQVRSIGADLANLYAQQV--GVPAVSFTLLDKNT-----GSGS 446
DB	339 GTMGNAAPQRIYAQLGGVYRTLSSTL---YRPFNIGINNQQLSVLDGTPEAYGTSSN 395
QY	447 VGGFTYKPFHTMQVCTQNYNTIDIEIPPENE--PLSRGYSHRLSHITSY--SPSKNASSP 502
DB	396 LPSAVYRKSGTV-----DSLDEIPQNNNVPPRGFSHRLSHVSMFRSGFNSVSI 447
QY	503 ARYGNLVPWATHRSADVNTVYSDKTIQIPVKAHTLVSGTTVYKGGPFGGHLKRTS 562
DB	448 IR---AMPFSWIHRSAFPNNIIPSSQITQIPLTKSTNLGSGTSVYVKGPGFGTGDILRRTS 504
QY	563 SGPLAYTSVSKSPLSQRYRARIYVASTTNLRLFTVTSGRTRYISINVNKTWNKGGDILTFN 622
DB	505 PGQISTURVNITAPLSQRYRARIYASTTNLQFTSIDGRPINQGNFSATMSSGNSLQSG 564
QY	623 TFDLATICTAFTPNYSDSLTVGADSPASGGEVTVDFKELLIPVNATFEABEDLVAKKAV 682
DB	565 SFRTVGGTTTPNFSGNSVFTLSAHVFNSGNEVIDRIEFVPAEVTFEAYDLERAQKAV 624
QY	683 NGLPFSKKD--ALQTSVTDYQVNOANLVECLSDIELYNPEKRMMDVAVKRAKRLVOARMLL 741
DB	625 NELFTSSNQIGLKTVDVTDYHIDQVSNLVECLSDSEFCLEKKEKSEKVKHAKRLSDERNLL 684
QY	742 QDTCFNRING--ENGWTCSGTIEVAEGDVLFPKORSRLTARSIEDTETVPTLYLQQIDDES 799
DB	685 QDPNLRGINQLDRGWRGSDTITIQGGDDVPFKNYVTLTGTFD---ECYPTLYLQKIDDES 741
QY	800 LLKPYTRYKLKGFSSQDLLEIKLIRHANOIVKNVPDNLPLDVLVFNSSCGIDRCSSBQQ 859
DB	742 KLKAYTRYQLGYIEDSQDLLEIYLIRYNAKHETVNVPGT--GSLWPLSAPSPIGKCAHH- 798
QY	860 YVDANTALENNGENGNSSSHAFSFHIDTCEIDLNENTGIWVFPKIPITTINGYATLGNLE 919

InterPro: IPR005639; endotoxin N.
 InterPro: IPR008979; Gal bind like.
 Pfam: PF00555; endotoxin; 1.
 Pfam: PF03944; endotoxin C; 1.
 Pfam: PF03945; endotoxin N; 1.
 SQ SEQUENCE 1180 AA; 133489 MW; 11998A6D1DCB62D CRC64;

Query Match 33.9%; Score 2050; DB 2; Length 1180;
 Best Local Similarity 39.3%; Pred. No. 3e-108;
 Matches 481; Conservative 174; Mismatches 414; Indels 156; Gaps 26;

QY 53 NPRL-----PGNPET-PISSSTVQTG---IGI---VGVQLGALGVPPAGQIASFYSP 97
 DB 4 NPNINECIPYCNLSNPVEVLGGRIETGTPIDISLSLTQFLSEFPVPGAGFVLGLVDI 63
 QY 98 IVGQLPSPSTVSVMEMIMKQVEDLIDQKITSVRKKTALAGLQGLDGLDVKYQSLKNWLE 157
 DB 64 IWGIFGPSQ---WDAFLVQIEQLINQRIEFPARNQAIISRLGSLNYQIYASFRWEA 119
 QY 158 NRNDTRARSVVVTVQYIALHLDLFVAKIPSPAIISQGVPLLSVYAQAANHLHLLLRDASIFG 217
 DB 120 DPTNPALREBMRIOFNMDNSALATTAIPLPAVQNYQVPLLSVYQAANHLHSLVRDVSFVG 179
 QY 218 AEWGPTPGBISTFYDQVTRTAQYSDYCVKWYNTGLDKLGTNAASWLKYHOFREMTLL 277
 DB 180 QRWGFAAATINSRYNDLTRIGNYTDYAVRWYNTGLERVWGPSRDMVRVYNQFRELTLT 239
 QY 278 VLQDLVLPYNDTRTYPIETTAOLTRVYTDPIVFNRETS--GGFCRWSLNSDISFSEVE 336
 DB 240 VLQDLVLPYNDTRTYPIETTAOLTRVYTDPIVFNRETS--GGFCRWSLNSDISFSEVE 290
 QY 337 SAVIRSPHLPDILSIEFYFT--TRAGLPLANNTEYLEYVWVGHSHK-----YKNTWASSALER 390
 DB 291 ---IROPHLMDILNRIITYDVRHG-----FNWMSGHQITASVPVPGSPGFAPPLFG 339
 QY 391 NYGTTISNKIKYDLANKDIQVRS-----LGLDANYAQQVGVVPYASFTLLDKNT 442
 DB 340 NAGNAAPPVL---VSLTGLGIFRTLSPLRYRRIILGSGPNQQLFVLDTGTFSPASLTNNL 397
 QY 443 GSGSVGGFTSKPHHTMOVCTQNTYDIDETPEPNE--PLSRGYSRLRSHLSTYSFSPKNAS 500
 DB 398 PS-----TIYRGTV-----DSLVDTPQDQNSVPPRAGFSRLRSHLSTYMLSQAGAV 444
 QY 501 SPARYGNLPFAWTHRSADVNTVYSDKITQIPVKAHTLVSGTTVIKPGFTGGMILKR 560
 DB 445 YTLR---APTFSWQHRAEENNIIPSSQITQIPLTSTNLGSGTSVVKVPGFTGDIILR 501
 QY 561 TSSGFLAYTSVSVKSPISQRYRARIYASTNNLFLFTTISGTRIYSINVKNTWNGKDDLT 620
 DB 502 TSPQISTLRVNTAPLSQRYRARIYASTNNLFLFTTISGTRIYSINVKNTWNGKDDLT 561
 QY 621 PNTFDLATICTAPTFSNVSDSLTVGADSPASGGEVYVDKPELIPVNAATFEABEDLOVAKK 680
 DB 562 SGFSRTVGTFTTTPNFNSGSSVFTLSAHVFNSGNEVYIDRIEFPVAVTFEAYDLEPRAQK 621
 QY 681 AVNGLFTSKKD--ALQTSVTVQVNOAANLVECLSDLELYPNKQMLMDVAKEAKRLVQARN 739
 DB 622 AVNELFTSSNQI GLKTDVTDYHIDQVSNLVECLSDLELYPNKQMLMDVAKEAKRLVQARN 681
 QY 740 LLQDTGPNRR--ENGWTSSTGLEVAGDVLPKRSRLTTSARBIDTETPTPYLVQOIID 797
 DB 682 LLQDPNFRGINRQLDRGWRGSDTIIOGGDDVFKENYVTLIGTFD---ECYPTVLYQKID 738
 QY 798 ESLLKPTVRYKLGKFGSSODLETKLIRHRAQIVKXNPDN---LLPDLVLPVNSCGGID 853
 DB 739 ESLLKPTVRYKLGKFGSSODLETKLIRHRAQIVKXNPDN---LLPDLVLPVNSCGGID 798
 QY 854 RCSEQQVVDANLALENNGENG--NMSDSHAFSPHIDTGEIDLMENTGIWVVFKIPTNGY 912
 DB 799 RCA--PHLEWNPDLDCRGEKCAHSHFSLDIDVGCTDMLNEDLGVWIFPKITQDGH 856
 QY 913 ATLGNLLEVBEGPLSGFTLRAQQEQOQWQDMRKARKGASEKAYTAARQAIDRLFADYQD 972

RESULT 8
 Q7WZT9
 ID Q7WZT9 PRELIMINARY; PRT; 1176 AA.
 AC Q7WZT9
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Delta-endotoxin.
 GN CRYLAA.
 OS Bacillus thuringiensis.
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 ON NCBI_TaxID=1428;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ly30;
 RA Yao J., Zhang J., Song F., Chen Z., Li C., Huang D.;
 RT "A novel cryIIAa gene from Bt strain ly30."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF384211; AAP80146.1;
 SQ SEQUENCE 1176 AA; 133169 MW; 725C14EA940B869B CRC64;

Query Match 33.9%; Score 2048; DB 2; Length 1176;
 Best Local Similarity 39.3%; Pred. No. 3.8e-108;
 Matches 480; Conservative 174; Mismatches 414; Indels 152; Gaps 26;

QY 53 NPRL-----PGNPET-PISSSTVQTG---IGI---VGVQLGALGVPPAGQIASFYSP 97
 DB 4 NPNINECIPYCNLSNPVEVLGGRIETGTPIDISLSLTQFLSEFPVPGAGFVLGLVDI 63
 QY 98 IVGQLPSPSTVSVMEMIMKQVEDLIDQKITSVRKKTALAGLQGLDGLDVKYQSLKNWLE 157
 DB 64 IWGIFGPSQ---WDAFLVQIEQLINQRIEFPARNQAIISRLGSLNYQIYASFRWEA 119
 QY 158 NRNDTRARSVVVTVQYIALHLDLFVAKIPSPAIISQGVPLLSVYAQAANHLHLLLRDASIFG 217
 DB 120 DPTNPALREBMRIOFNMDNSALATTAIPLPAVQNYQVPLLSVYQAANHLHSLVRDVSFVG 179
 QY 218 AEWGPTPGBISTFYDQVTRTAQYSDYCVKWYNTGLDKLGTNAASWLKYHOFREMTLL 277
 DB 180 QRWGFAAATINSRYNDLTRIGNYTDYAVRWYNTGLERVWGPSRDMVRVYNQFRELTLT 239
 QY 278 VLQDLVLPYNDTRTYPIETTAOLTRVYTDPIVFNRETS--GGFCRWSLNSDISFSEVE 336
 DB 240 VLQDLVLPYNDTRTYPIETTAOLTRVYTDPIVFNRETS--GGFCRWSLNSDISFSEVE 290
 QY 337 SAVIRSPHLPDILSIEFYFT--TRAGLPLANNTEYLEYVWVGHSHK-----YKNTWASSALER 390
 DB 291 ---IROPHLMDILNRIITYDVRHG-----FNWMSGHQITASVPVPGSPGFAPPLFG 339
 QY 391 NYGTTISNKIKYDLANKDIQVRS-----LGLDANYAQQVGVVPYASFTLLDKNT 442

Qy 621 ENTFLATICTAFTFSNYSDSLTVGADSPASGGVYVDVKFELIPVNATFEABEDLDVAKK 680
Db 562 SGSPRTVGTPTPNFNSGSSVFTLSAHVFNSEVNIIDRIEFVPAEVTFFBAFYDLERAQK 621
Qy 681 AVNGLFTSKKD-ALQTSVTDYQVNOAANLVECLSDLEYLPNEKRLMDAVKAKRLVQARN 739
Db 622 AVNELFTSSNQIGLKTVDVTHIDQVSNLVECLSDLEYLPNEKRLMDAVKAKRLVQARN 681
Qy 740 LLODTGNRING--ENGWTSSTGIEVAGDVLFDKRLSLTSAREIDTETPTLYQOID 797
Db 682 LLODPNFRGINRQLDRGWSGTDITIQGGDDVFKNVYVTLTGTFD---ECYPTLYQKID 738
Qy 798 ESLLKPYTRYKLGFGISSQDLLEIKLIRHRANOIVKRVNPDN---LLPDLVLPVNSCGGID 853
Db 739 ESLLKPYTRYKLGFGISSQDLLEIKLIRHRANOIVKRVNPDN---LLPDLVLPVNSCGGID 798
Qy 854 RCSEQQVVDANLALNNGENG-NMSSDSHAPSFIHDTGIEDLNENTGIWVFKIPTNGY 912
Db 799 RCA--PHLEWNPDLDCSCRDGEKCAHSHHPSLSDIDVGCTDLNHDLGWVWIFKIKTQDGH 856
Qy 913 ATIGNLELVEEGLSGETLERAOQOQWODKMARKGASKEYAAKQADRLFDYQD 972
Db 857 ARGNLEFLBEKPLVGEALARKVAKKWKREKLEWETNIVYKEAKESVDALFVNSQY 916
Qy 973 QKLSGVEMSDMLAAQNLVOSIPYVNDALPEIFGNYTSTFELTNLQOAWNLIDLNA 1032
Db 917 DQLOQDNTMIAMIHAADKRVHSIREAYLPVLSVPGVNAALFEELBEGRISTAFSLYDARNV 976
Qy 1033 IPNGDFRGLSDMNATSDVNV-QQLSDTVLVPVNSQVSOQFTVPQNTRYVLRVARK 1091
Db 977 IKNGDFNGLSCWNVKGVHVDVEQNNQSVLVPVNEAEVSEVVRVCPGEGYILRVATYK 1036
Qy 1092 EGVGDGVVIRDXGANTETLTFN-----ICDD----- 1118
Db 1037 EGYEGCVTHIEKNNTDELKFSVCVEEBIYPNNTVTCNDTIVNQEYGGAYTSRNGYN 1096
Qy 1119 DTGVLSDOTS-----YITKTVETPTPEQWID 1147
Db 1097 EAPSVADYASVYEKSYTDGRRENPCFENRKYRDYTPLPVGYVTKLEYFPETDKWIE 1156
Qy 1148 MSETGVPNTIESVELVEE 1167
Db 1157 IGETEGTFIVDSVELLME 1176

RESULT 10
Q9RC30
ID Q9RC30 PRELIMINARY; PRT; 1176 AA.
AC Q9RC30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 135 kDa insecticidal protein.
OS Bacillus thuringiensis (subsp. kurstaki).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD-1-02;
RA Hou B.K., Chen Z.H.;
RL EMBL; AF154676; AAD55382.1; -
DR HSP; P02965; ICIY.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; Endotoxin C.
DR InterPro; IPR005639; Endotoxin N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin C; 1.
DR Pfam; PF03945; endotoxin N; 1.

SO SEQUENCE 1176 AA; 133010 MW; FCB069D0B081D8C4 CRC64;
Query Match 33.7%; Score 2034; DB 2; Length 1176;
Best Local Similarity 39.2%; Pred. No. 2.4e-107;
Matches 478; Conservative 173; Mismatches 417; Indels 152; Gaps 26;
Qy 53 NPDL-----FCNPET-FISSSTVQVG-----IGI-----VGQVLGALGVPVPAQQLASFYSF 97
Db 4 NPNIECIPYNCLSNPEVEVLGGRIETGYTPIDISLSTQFLSSEFVGAGFVLGLVDI 63
Qy 98 IVGQLWPSSTVSVWEMTMKQVEDLIDOKITDSVRKTAGLQGLGDLGVYOKSLKNWLE 157
Db 64 IWGIFGFSQ-----WDTLVQIEQLNQIBEFARNOALSRLEGSLNLQIYABSFREMA 119
Qy 158 NRNDTRARSVVVTVYIALELDFVAKIPSAISFALSGOEVPLLSVYQAANLHLLLRDASIFG 217
Db 120 DPTNPALREEMRIQFNDMNSALTTAIPLLAVQNVQVPLLSVYQAANLHLSVLVDVSVFG 179
Qy 218 AEWGFTPEEISTFVDQRVTRTAQYSDYCVKQWNTGLDKLKTNAASWLKYHQPRENTLL 277
Db 180 QRMGFDAAATNSRYNDLTRLIGNYTDVAVRWYNTGLERVMGPDSDRWVRYNQFREELTIT 239
Qy 278 VLDLVALFPNVDTTYPIETTAQLTRVYTDPIVFNRETS--GGPCRRNSLNSDISFSEVE 336
Db 240 VLDLVALFPNSYDSRYPIRTVSQLTRVYTNVLENFDGSPFRGMAQRIEQN----- 290
Qy 337 SAVIRSPLHFDLSEISFYT--TRAGLPLNNTVLEYLVWGHSHIK-----YKNTNASSALER 390
Db 291 ---IRQPHLMDLINSIITYDVRG-----FNWMSGHQITASPVGSPGPEAPPLFG 339
Qy 391 NYGFTSNKI KYVDLANKDIPQVRS-----LGADLANYTAQVGVYPYASFLLDKRT 442
Db 340 MAGNAAPPVL--VSLTGLGIFRTLSSPLRYRRIILGSGPNQBELFVLDGTFEFSFALTTNL 397
Qy 443 GSGSVGGFTYKPHHTMQVCTQNTYNTIDEIPENE--PLSRGYSHLSHITSYSYSKNAS 500
Db 398 PS-----TIYRQGTV-----DSLDPVPPQNSVPPRAGFSHRLGHVTHLSQAGAV 444
Qy 501 SPARYGNLPFAWTHRSADVNTVYSDKITQIPVVKAHTLVSGTTVVKGPFTGGNILKR 560
Db 445 YTLR--APTFSMQHRSABFNIIIPSSQITQIPLTKSTNLGSGTSVVKGPGTGGDILRR 501
Qy 561 TSSGPLAYTSVSKSPLSQRYRARIYASTNBLFVTISGTRIYSINVKTNMGDDLT 620
Db 502 TSPQISTLRVNTAPLSQRYRARIYASTNBLFVTISGTRIYSINVKTNMGDDLT 561
Qy 621 ENTFLATICTAFTFSNYSDSLTVGADSPASGGVYVDVKFELIPVNATFEABEDLDVAKK 680
Db 562 SGSPRTVGTPTPNFNSGSSVFTLSAHVFNSEVNIIDRIEFVPAEVTFFBAFYDLERAQK 621
Qy 681 AVNGLFTSKKD-ALQTSVTDYQVNOAANLVECLSDLEYLPNEKRLMDAVKAKRLVQARN 739
Db 622 AVNELFTSSNQIGLKTVDVTHIDQVSNLVECLSDLEYLPNEKRLMDAVKAKRLVQARN 681
Qy 740 LLODTGNRING--ENGWTSSTGIEVAGDVLFDKRLSLTSAREIDTETPTLYQOID 797
Db 682 LLODPNFRGINRQLDRGWSGTDITIQGGDDVFKNVYVTLTGTFD---ECYPTLYQKID 738
Qy 798 ESLLKPYTRYKLGFGISSQDLLEIKLIRHRANOIVKRVNPDN---LLPDLVLPVNSCGGID 853
Db 739 ESLLKPYTRYKLGFGISSQDLLEIKLIRHRANOIVKRVNPDN---LLPDLVLPVNSCGGID 798
Qy 854 RCSEQQVVDANLALNNGENG-NMSSDSHAPSFIHDTGIEDLNENTGIWVFKIPTNGY 912
Db 799 RCA--PHLEWNPDLDCSCRDGEKCAHSHHPSLSDIDVGCTDLNHDLGWVWIFKIKTQDGH 856
Qy 913 ATIGNLELVEEGLSGETLERAOQOQWODKMARKGASKEYAAKQADRLFDYQD 972
Db 857 ARGNLEFLBEKPLVGEALARKVAKKWKREKLEWETNIVYKEAKESVDALFVNSQY 916
Qy 973 QKLSGVEMSDMLAAQNLVOSIPYVNDALPEIFGNYTSTFELTNLQOAWNLIDLNA 1032
Db 917 DQLOQDNTMIAMIHAADKRVHSIREAYLPVLSVPGVNAALFEELBEGRISTAFSLYDARNV 976

```
QY 1033 IPNGDFRNLGSDNATSDVNV-QQLSDTSVLVLPNNNSOVSOQFTVQPNRYVLRVARTAK 1091
DB 977 IKNGDFNGLSCWVKGHVDVEQNNQRSLVVPWEAEVSOEVRVCPGEGVILRVATYK 1036
QY 1092 EGVGGVYIIRGANOTETLTN-----ICDD-----1118
DB 1037 EGTGGCVTHIEHNITDELKPSNCVEEIIYNNVTTCNDYTNQBEYGGVATSNRGN 1096
QY 1119 DTGVLSDQTS-----YITKVFETPSTEQWID 1147
DB 1097 EAPSPADYASYVEEKSYTDGRNCPENRGYRDYTPLPVGVYKLEVPETDKWIE 1156
QY 1148 MSTEGVNIESVELVLEE 1167
DB 1157 IGETGTFTVDSVELLME 1176

RESULT 11
Q93T21
AC Q93T21; PRELIMINARY; PRT; 1155 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Crystal protein CryIAB16.
GN CRYIAB16.
OS Bacillus thuringiensis (subsp. israelensis).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1430;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=407;
RA Yu J., Tan L., Wu D., Pang Y.;
RT "Molecular characterization of a silent gene encoding a 130-kilodalton
RT crystal protein from Bacillus thuringiensis subsp. israelensis.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375608; AAK5546.1; -.
DR PIR; A29125; A29125.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF00555; endotoxin; Gal bind_like.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
SQ SEQUENCE 1155 AA; 130747 MW; 7F0C98E0100C7698 CRC64;

Query Match 33.6%; Score 2032.5; DB 2; Length 1155;
Best Local Similarity 38.9%; Pred. No. 2.9e-107;
Matches 473; Conservative 185; Mismatches 392; Indels 167; Gaps 29;

QY 53 NPEL-----FGNPET-FISSVTQTG---IGI---VGQVIGALGVFPAGQIASFYSF 97
DB 4 NPNINECIPYNCLSNEPEVLGGERIETGYTPTIDISLSLTQFLSEFVFGAGFVLVDI 63
QY 98 IVGQLWPSSTVSWEIMKQVEDLDOKLTDVSRKTAGLAGLOGLDVYOKSLKNWLE 157
DB 64 IWGIFGPSQ-----WDALVQIQLINQRIEFPARNQAIISRLGSLNLYQIYAESFWEA 119
QY 158 NENDTRARSVVVQYIALBELDFAKIPSPAISQEVPLLSVYAQAANLHLLLRASIFG 217
DB 120 DPTNPALREEMRIQFNWNSALTATPLPAVQNYRVPLLSVYVQAVNLHLSVLRLVDFG 179
QY 218 AEWGTFGISITPDRQVTRTAQSYCYKWNVTGLDKLGTNAASLWKYHQRFRMTLL 277
DB 180 QRWGFDAAATINSRYNDLTRIGNYTHAVRWYNTGLRVWGPDSRDWIRYQFRRELTLT 239
QY 278 VLDLVALPNYDTRTPYIETTAQLTEVYTDPIVENRSTSGGFCRWLSLNSDISFSEVBS 337
DB 240 VLDIVSLFPNYSRTYPIETTSQLTREIYTNPVLENFD--GSF--RGSAGQ-----IRG 289
```

```
QY 338 AVIRSFPLEDILSEIEFTTTRAGLPLNNTEYLYEYVGHSHK-----YKNTNASSALERNY 392
DB 290 S-IRSFPLMDILNSITTYT-----DAHRGEY--YMSGHQIMASPVGSGPEFTFPL---Y 338
QY 393 GTITSNKIKYDLANKDIFQVRSIGADLANYYAQVY--GVFYASFTLLDKNT---GSGS 446
DB 339 GTMGNAAPQORIVAQGGQVYRLSTL---YRRPFIIGNNQQLSVLDOTFEAYGTSSN 395
QY 447 VGGFTYSKPHHTMQVCTQNYNTIDEIPPEB--PLSRGYSHRLSHITSY--SPSKNASSP 502
DB 396 LPSAVTRKSGTV-----DSLDRIPPQNNVPRQGFSHRLSHVSMPSGFSNSVSI 447
QY 503 ARYGNLPVPAWTHRSADVNTVYSKIQIOPVKAHTLVSGTIVVIGPGPTGNGILKRTS 562
DB 448 IR---APMFSWTHRSABFNIIIPSSQITQIPLTKSTNLGSGTSVVKVPGFTGDDILRST 504
QY 563 SGPLAVTSVSKSPLSQRYRARIYASTTNLRLFTVITSGTRIVSYINVKTMNKGDDLTN 622
DB 505 PQQISTLRVNTITAPLSQRYRARIYASTTNLQFTSIDGRPINQGNFSAATMSGSNLQSG 564
QY 623 TFDLATIGTAFPTFSNYSDSLTVGADSFASGGEVYVDKFLIPVNNATFEAEEDLDVAKAV 682
DB 565 SFTVGTPTTFFNPSGSSVPTLSAHVFNSENVYIDRIEFPVPAEVTFEAYDLERAQKAV 624
QY 683 NGLFTSKKD-ALQTSVTDYQVNOAAMLVESCLSDLYPNBKRMLWDVAKRLVQARNLL 741
DB 625 NELFTSSNQIGLKTQDVTYHIDQVSNLVECLSDFCLEKELSKBVKHAKRLSDERNLL 684
QY 742 QDTGNNRING--ENGWGTSGTIEVARGDVLFKDRSLRLTSAREIDTETPTLYVQQLDES 799
DB 685 QDPNFRGINQLDRGWRGSTDITIQGGDDVFKENYVTLTGTFD---ECYPTLYIKLDES 741
QY 800 LKPYTRYLKGFGIGSSQDLKLRHRANQIVKNNVPDMLLPDLPVNSCGGIDRCSEBQ 859
DB 742 KLAYTRYQLRGYIEDSDLEIYLYRYNAKHETVNVPGT--GSLWPLSAPSPIGKCAHH- 798
QY 860 YVDANLALNNGENGNMSSDSHAPSFIHDTGEIDLNNENTGIWVVKIPITNGYATLGNLE 919
DB 799 -----SHHFSLDIDVGCTDLNEDLGVVWVIFKIKTQDGHARLGNLE 838
QY 920 LYEBGLSGETLERAQOQOQODKWKARKGSEKAYAAKOAIIDLFLADYQDQKLSGV 979
DB 839 FLEEKPLVGEALARKVRAEKRRDKRLEWETNIVYKKAESVDVAVFVNSQVDRLOQDT 898
QY 980 ENSDMLAQNLYQSIPYVYNDALPEIPGNNYTSFTLTNRLQOANLYDLRNAIPNGDPR 1039
DB 899 NIAMIHADKRVHSIREAYLPESLVIPGVYNAALFELEGRIPTAFSLYDARNVKNKDPN 958
QY 1040 NGLSDNATSDVNV-QQLSDTSVLVLPNNNSOVSOQFTVQPNRYVLRVARTEGVDGY 1098
DB 959 NGLSCWNVKGVHDVEEQNNHRSVLVVPWEAEVSOEVRVCPGEGVILRVATYKBEYBGC 1018
QY 1099 VLIIRGANOTETLTN-----ICDDTGV-----1122
DB 1019 VTIIHEIENNTABLKFSNCVEEIVYNNVTVCNDYATQBEYGTYSRIRGYGAYESNS 1078
QY 1123 -LSADQTS-----YITKVFETPSTEQWIDMSE 1150
DB 1079 SVPADYASAYEAKYTDGRDNPNCESNRGVDYTPLPAGVTVRELEVPPETDKWIEIGE 1138
QY 1151 TEGVNIESVELVLEE 1167
DB 1139 TEGTFTVDSVELLME 1155

RESULT 12
O06894
AC O06894; PRELIMINARY; PRT; 1171 AA.
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

CryIEA4.	
GN CryIEA4.	
OS Bacillus thuringiensis.	
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.	
NCBI_TaxID=1428;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=LBIT-147;	
RA Barboza-Corona J.E., Lopez-Meza J.E., Ibarra J.E.;	
RT "Cloning and expression of the cryIIIA4 gene of Bacillus thuringiensis	
RL and the comparative toxicity of its gene product.";	
RR World J. Microbiol. Biotechnol. 14:437-441(1998).	
RN [2]	
RP SEQUENCE FROM N.A.	
RC STRAIN=LBIT-147;	
RA Ibarra J.E., Barboza-Corona J.E.;	
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.	
EMBL; U94323, AAD04732.1; -.	
DR PIR; A37829; A37829.	
DR HSP; P02965; LCYI.	
DR GO; GO:0015070; P:toxin activity; IEA.	
DR CO; GO:0006952; P:defense response; IEA.	
DR InterPro; IPR001178; Endotoxin.	
DR InterPro; IPR005638; endotoxin_C.	
DR InterPro; IPR005639; endotoxin_N.	
DR pfam; PF00555; endotoxin; 1.	
DR pfam; PF03944; endotoxin_C; 1.	
DR pfam; PF03945; endotoxin_N; 1.	
SQ SEQUENCE 1171 AA; 133310 MW; 333603MIA782523A CRC64;	

Query Match 33.4%; Score 2016.5; DB 2; Length 1171;
 Best Local Similarity 38.3%; Pred. No. 2.4e-106;
 Matches 464; Conservative 176; Mismatches 412; Indels 159; Gaps 22;

Qy	59	NPETFI-----SSSTVGTGIGVQLGALGVFPAGQIASFYFIVGLWPSSVTSVWEM	113
Dd	18	NPENEILDIERNSTVATNIAREISRLASATPIGGILLGLDFDAIWGSIPESQ----	WDL 73
Qy	114	IMKQVEDLDOKITDSVRKTALAGLGDLGVYQKSLKNLENRNDTRARSVVVITY	173
Dd	74	FLEQIELLDOKILEEPARNOAISREGLISSLYGIYTEAFREWADPTNPALKEEWRTPN	133
Qy	174	ALEDLPFAKIPSPAISGOEVPLLSVYAQAANHLHLLLRDASIPGAENGFTPGRISTFYDR	233
Dd	134	DMNGILVTAIPLFSVQNYQVPFLSYVYAQAANHLSQLRDVSVFQAWGFDIATINSRYND	193
Qy	234	QVTTAQVSDYCVKWYNTGLDKLKGTNA--ASWLKYHQFRREMTLLVLDLVALFNVDYTRT	292
Dd	194	LTRLPIPIYTDVAHWYNTGLDLPRTGGLRNWARPNQFRRLTIISLVLDIISFFRNYSRL	253
Qy	293	YPETTAQLTREVTDPDTVEENRETSGGFCRWLSNLSDISPSEVESAVTRSPLHPDLISEI	352
Dd	254	YPITPSQTREVTDPVNITDTRVG-----ESPENIENSALRSPLMDPLNVL 303	
Qy	353	EFTYTRAGLPLNNTEYLEFWVGHISKYKNTNASSALER-NYG-----TITS 397	
Dd	304	TIDT-----DLRGVHYWAGRHTVSHFTGSSQVLTTPQXGNSQNASPRERTIAPSTPPG	356
Qy	398	NKIYYDLANKDIFQVRSLGADLANYYAQVGVPPYASTLLDRNKTSGSGVGFYSGPHT 457	
Dd	357	LNLFYRTLNSNPFRRSENITPTL-----GINNVQGVGFIOFPNN 394	
Qy	458	TMQVCTQNYTID---BIPPENELSRGYSHRLSHI---TSYSPSKNASSPARYGNLPPVA 512	
Dd	395	A-EVLVYRSGTVDSLNELPIPDGENSLVGYSHRLSHVTLTRSLYNINITS-----LPTFV 447	
Qy	513	WTHSADVNTVYSDKITQIPWAHTLI VSGTWIKGPFGTCGNILKRTSSCPLAYTSVS 572	
Dd	448	WTHSATNTNTINPDIITQIPLVKGFRLCGGTSVILKGPFGTGGDILLRNTTIGEVSVLQVN 507	
Qy	573	VKSPLSQRYSRARIVASYATTNLRLPVITSGTRIYSINVNKTWNKGDLLTFNTFDLATICTA 632	

Db 305 NNLTIFDWFVSGRN-----PYWGCHRVISSLLGGNITSPYGRANQBPSPFTNGP 359
Qy 410 IFQVRSAGADLANYYAQQVGYVYASFTLLDKNTGSGSGVGGFTYKPHMTMQVCTQNTYNI 469
Db 360 VP--RTLSPNTRLRLLQPPWAPPENLGRVGVGFSTPTNSFTYGRGTG-----DSL 409
Qy 470 DEIPPENE--PLSRGYSHRLSHITSYSPKSNASSPARYGNLPVPAWTHRSADVNTVYS 527
Db 410 TELPPEQNSVPPREGYSHRLCHAT---FVQSGTFFLTGTV-VFSWTHRSATLNTIDPE 465
Qy 528 KITQIPVVKAHVLTGVTGKGTGKTLSSGFLAYTSVSKSPLSQRIRARY 587
Db 466 RINQIPLVKGRWGGISVITGPGFTGDLIRNFTGDFVSLQVINSPIQRLRFRY 525
Qy 588 ASTTNRLFLVT-----ISGTRIYSNVNKNMKGDDLTNTFLATIGTAFTFSYSD 640
Db 526 ASSRDARVIVLTGAASGVGVQSVNMPLOKMTBEIGENLTSTRTFYDPSNPPSPFRAND 585
Qy 641 SLTV-----GADSFASGGEVYVDPKPHLPVATFEABEDLDVAKKAVNGLFTSKD-AL 693
Db 586 IIGISERPLFGAGISS-GELYIDKIRIILADATFEABESDLERAQKAVNALFTSSNQIGL 644
Qy 694 QTSVTDYVQVQNAANLVELCLSDLYPNKSKMLWDVAKKAKLVQARNLLQDTGENIN--G 751
Db 645 KTDVTDYHIDQVSNLVDCLSDDEFCLDEKSEKSEKVKAKRSLDERNLLQDPNFRGINRQP 704
Qy 752 ENGHTSGTGEVAREGDLVLFKDRSLRLTSARIDTETPTLYQOIDSLLKPYRYKLKG 811
Db 705 DRWGGSTDTIQGDDVFKENVYTLPGT--VD-ECYPTLYQKIDSKLKAYTRYBLRG 761
Qy 812 FIGSQOLEKLIHRANQIVNVPDN-----LLPDVLPVNSCGGIDRCSEQQYVDANLAL 867
Db 762 YIEDSQLEIYLIIRYNAKHEIVNVPGTGSLWPLSAQSPIGKCGBNRCA--PHLEWNPDL 819
Qy 868 ENNGENG-NMSSDSHAFSPHIDTGEIDLNENTGIVVVKIPTTNGYATLGNLELVBERGL 926
Db 820 DCSRDGEKCAHSHHFTLDIDVGCTDLNEDLGWVVFVKITQDGHAKGLNLEFLEKPL 879
Qy 927 SGETLEAAQQOQOQWQDMARKGASBKAYAAQAIIDRLPADYQOQKINSGVMSDLA 986
Db 880 LGEALAVKRAEKWKREKQLETNIVYKEAKESVDALPVNSQYDRLQVDTNMIHA 939
Qy 987 AQNLVQSPYVYNALPEIPGANTSTFELTNRLQQAANLVDLRLNAPNGDFRGLSDWN 1046
Db 940 ADKRVHRIEAYLPESLVPQVNAAIPEELRGRIFTAYSLYDARNVIRKNGDFNGLLCWN 999
Qy 1047 ATSDVNV-QQLSDTSVLVIPNNNSQVQFTVQPNRYVRLVATKRGVGDGVYVIRGGA 1105
Db 1000 VKGHVDVEQNNHRSVLVIPLEAEVSOEVRVCPGRGYILRVATYKBYGEGCVTIHIE 1059
Qy 1106 NOTELTFN-----ICDDDTGV-----LSADQT 1128
Db 1060 DNTDELAFSCVBEVYPNNTVTCNNYTGTOBEYGYTSSENQGYDRAFGNPNVPADYA 1119
Qy 1129 S-----YITKVTFTPTSPTEQWIDMSFEGVNI 1157
Db 1120 SVYEKSQYDGRNCPESNRGQYDTPLPAGYVTKOLEYFPETDKVWIEIGETGFIV 1179
Qy 1158 ESVELVLEEE 1167
Db 1180 DSVELVLEEE 1189

RESULT 15
Q45749
ID Q45749 PRELIMINARY; PRT; 1174 AA.
AC Q45749;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Crystal protein (Crystal delta-endotoxin).
GN CRY1PB.
OS *Bacillus thuringiensis*.

Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
NCBI_TaxID=1428;
OC [1]
RN SEQUENCE FROM N.A.
RC STRAIN=BTS00349A;
RA Lambert B.;
RT "NO INFORMATION";
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
RP [2]
RN SEQUENCE FROM N.A.
RC STRAIN=B-Pr-88;
RA Li C., Zhang J., Huang D., Li G.;
RT "A crystal endotoxin from *Bacillus thuringiensis* strain B-Pr-88";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z22512; CAAB0235.1; -;
DR EMBL; AF336114; AAO13295.1; -;
DR PIR; S32649; S32649.
DR HSSP; P02965; ICIV.
DR GO; GO:0015070; P:toxin activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin C.
DR InterPro; IPR005639; endotoxin N.
DR InterPro; IPR008979; Gal_bind_Like.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin C; 1.
DR Pfam; PF03945; endotoxin N; 1.
SQ SEQUENCE 1174 AA; 133350 MM; 8C7F123F9446P15C CRC64;
Query Match 32.9%; Score 1989.5; DB 2; Length 1174;
Best Local Similarity 38.0%; Pred. No. 8.4e-105; Indels 143; Gaps 25;
Matches 458; Conservative 191; Mismatches 413;
Qy 59 NPRTFISSTVQTG-----IGIVGVGLGALGVPPAGQIASFYFIVGQLWPSSTVSVW 111
Db 17 NPVEYLLSEESTGRLPLDISLSLTRLFLSFP-VPGVGVAFLGFLINGFITPSE---W 71
Qy 112 EMIMKQVEDLIDQKITDSVRKTALAGLQGLDGVYQKSLKNWLENDRNDRASVVVWQ 171
Db 72 SLFLQIQEQLIEQRIETLERNRAITTLGLADSVEYLEALREBENPNAQLREBVRIR 131
Qy 172 YIALELDPVAKIPSPFAISGQEVPLLSVYQAANLHLLLRDASIFGASWGFPGHISTFY 231
Db 132 FANTDDALITAINNLTITFEIPLLSVTQAANLHLSLRDAVSFGQGWGLDIATVNNHY 191
Qy 232 DRQVTRTAQYSDYCVKWTNTGLDKLKGNTAASWLKYHOFREMTLLVLDLVALFPNYDTR 291
Db 192 NLLNLHRYTEHCLDVTYNOGLENLGNTTQWRSFNQFRRELTLTLDIVLVALFPNYDAR 251
Qy 292 TYPIETTAQLPRRVYTPDIVFNRETSGGFCRRMSLNSDISFSEVESAVIRSPHLPDILSE 351
Db 252 AYPQTSSQLTRRYTSSVIEDSPVSA-----NIPNGFNRAEFGV-RPPLHADMNS 302
Qy 352 IEFYTRAGLPLANNTEYLEVYVGHISIKYKNTNASSALERNYGTIT-----SNKIKY 402
Db 303 L-FVTA-----FTVRSQTWGGHLSVSRNTAGNIPPIYGFIPNGGAIWATDSDPRPF 355
Qy 403 YDLANKDIFQVRSIGADLANYYAQQVGYVYASFTLLDKNTGSGSGVGGFTYKPHMTMQVC 462
Db 356 YRTLSDPVP-VRG-GFGNPHVVLGLGVAF-----QQTGTN-----HTRTPN 396
Qy 463 TQNTYNTIDEIPPENEPLS--RGYSHRLSHITSYSPKSNASSPARYGNLPVPAWTHRSADV 520
Db 397 SGTIDSLDEIPPDQNSGAPMNDYSHVLNHHVTFVRWPKGIAGSDSW-RAPMFSWTHRSADR 455
Qy 521 TNTVYSDKIQIPIVVKAHVLTGVTGKGTGKTLSSGFLAYTSVSKSPLSQR 580
Db 456 TNIINPNIIITQIPAKVAHNLHSGSTVVRGPGFTGDDLRLRTNTGTFADIRVNIITPLSOR 515
Qy 581 YRARIYASTNRLFLVFTISGTRIYSNVNKNMKGDDLTNTFLATIGTAFTFSYSD 640
Db 516 YRVIIRYASTTDIAOFFTRINGTSVNOGNFQRTMNRGNLSEGNFRTAGTSPFSPSNAQS 575

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:30:40 ; Search time 65 Seconds
(without alignments)
5072.814 Million cell updates/sec

Title: US-10-089-678-1

Perfect score: 6044

Sequence: 1 MSPNNQNEYILDASSTSV.....MSSETGVFNIESVHLVLEE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6044	100.0	1167	5	AAU80281
2	4008	66.3	1157	2	AAU80281
3	4007	66.3	1157	2	AAU80281
4	4007	66.3	1157	2	AAU80281
5	4007	66.3	1157	2	AAU80281
6	3995	66.1	1157	2	AAU80281
7	3994	66.1	1157	2	AAU80281
8	3662.5	60.6	1149	2	AAU80281
9	3662.5	60.6	1149	2	AAU80281
10	3662.5	60.6	1149	2	AAU80281
11	3246.5	53.7	1206	5	AAU9255
12	3152	52.2	1169	2	AAU80281
13	3128	51.8	1157	2	AAU80281
14	3128	51.8	1157	2	AAU80281
15	3126	51.7	1157	2	AAU80281
16	3125	51.7	1157	2	AAU80281
17	3125	51.7	1157	2	AAU80281
18	3124	51.7	1157	2	AAU80281
19	3124	51.7	1157	2	AAU80281
20	3124	51.7	1157	2	AAU80281
21	3124	51.7	1157	2	AAU80281
22	3124	51.7	1157	2	AAU80281
23	3124	51.7	1157	2	AAU80281
24	3123	51.7	1157	2	AAU80281
25	3123	51.7	1157	2	AAU80281

26	3123	51.7	1157	2	AAW84566	AAW84566 Amino aci
27	3123	51.7	1157	2	AAW84569	AAW84569 Amino aci
28	3123	51.7	1157	2	AAW84568	AAW84568 Amino aci
29	3123	51.7	1157	4	AAW82178	AAW82178 Cry9C #1.
30	3122	51.7	1157	2	AAW84583	AAW84583 Amino aci
31	3122	51.7	1157	2	AAW84592	AAW84592 Amino aci
32	3121	51.6	1157	2	AAW84572	AAW84572 Amino aci
33	3121	51.6	1157	2	AAW84577	AAW84577 Amino aci
34	3120	51.6	1157	2	AAW84590	AAW84590 Amino aci
35	3120	51.6	1157	2	AAW84579	AAW84579 Amino aci
36	3119	51.6	1157	2	AAW84589	AAW84589 Amino aci
37	3118	51.6	1157	2	AAW84587	AAW84587 Amino aci
38	3117	51.6	1157	2	AAW84573	AAW84573 Amino aci
39	3115	51.5	1157	2	AAW84567	AAW84567 Amino aci
40	3115	51.5	1157	2	AAW84585	AAW84585 Amino aci
41	3114	51.5	1157	2	AAW84580	AAW84580 Amino aci
42	3109	51.4	1157	2	AAW84574	AAW84574 Amino aci
43	3100.5	51.3	1156	2	AAW46857	AAW46857 Bacillus
44	3100.5	51.3	1156	2	AAW24960	AAW24960 Bacillus
45	3100.5	51.3	1156	4	AAU02034	AAU02034 B. thurin

ALIGNMENTS

RESULT 1
AAU80281
ID AAU80281 standard; protein; 1167 AA.
XX AAU80281;
AC
XX
DT 30-JUL-2002 (first entry)
XX
DE Bacillus thuringiensis insecticidal protein.
XX
XX Insecticide; transgenic; Coleoptera larvae.
XX
XX Bacillus thuringiensis.
OS JP2002045186-A.
PN
XX
PD 12-FEB-2002.
XX
PF 03-AUG-2000; 2000JP-00236140.
XX
PR 03-AUG-2000; 2000JP-00236140.
XX
PA (SDSB-) SDS BIOTECH CORP.
XX WPI; 2002-356468/39.
DR A protein having insecticidal activity, a DNA encoding said protein, and an agent and a method for preventing harmful organisms.
PT
XX
PS Claim 1; Page 9-12; 19pp; Japanese.
XX
XX This invention relates to a crystalline protein comprising a fully defined sequence and the nucleotide sequence encoding this protein. The protein of the invention is an agent for preventing harmful organisms comprising Bacillus thuringiensis serovar galleriae SDS502, its mutant or a microbe transformed by a DNA encoding the protein. This microbe can be used to produce a protein containing the protein, or containing a protein having insecticidal activity produced by the SDS502, its mutant or a transformed microbe, a microbe which is transformed by using the above DNA and produces the above protein having insecticidal activity, a plant or a seed transformed by using the above DNA, and Bacillus thuringiensis serovar galleriae SDS502 producing a protein comprising and producing a protein showing insecticidal activity. The protein of the invention may have insecticidal activity. The agent is used for preventing Coleoptera larvae. This sequence represents the Bacillus thuringiensis insecticide protein of the invention
XX Sequence 1167 AA;
SQ

```
Query Match      100.0%; Score 6044; DB 5; Length 1167;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPNNQNEVEIILDASSSTSVSDNSVRYPLANDQTTTLQNNMYKDYLRMSGEGNPELFGNP 60
DB 1 MSPNNQNEVEIILDASSSTSVSDNSVRYPLANDQTTTLQNNMYKDYLRMSGEGNPELFGNP 60

QY 61 ETPFISSTVQTGIGIVGVLGALGVPFAGQIASFYSTVGQWPSTSVSWEMIMKQVED 120
DB 61 ETPFISSTVQTGIGIVGVLGALGVPFAGQIASFYSTVGQWPSTSVSWEMIMKQVED 120

QY 121 LIDQKITDSVRKTAAGLGLGDLVDVYQKSLKNWLENRNDTRARSVVVTVYIALLELDFV 180
DB 121 LIDQKITDSVRKTAAGLGLGDLVDVYQKSLKNWLENRNDTRARSVVVTVYIALLELDFV 180

QY 181 AKIPSPAISSQGVPLLSVYQAQANLHLILRRDASIPGAEWGFTGCEISTFYDRQVTRTAQ 240
DB 181 AKIPSPAISSQGVPLLSVYQAQANLHLILRRDASIPGAEWGFTGCEISTFYDRQVTRTAQ 240

QY 241 YSDYCVKMYNTGLDKLKGNTNAASWLKYHQFREMTLLVLDLVALPNDYDTRTYPIETTAQ 300
DB 241 YSDYCVKMYNTGLDKLKGNTNAASWLKYHQFREMTLLVLDLVALPNDYDTRTYPIETTAQ 300

QY 301 LTRVYVTDPIVNRRTSGGFCRRWSLNSDISPSEVESAVIRSPHFLDILSEIEBFTYTRAG 360
DB 301 LTRVYVTDPIVNRRTSGGFCRRWSLNSDISPSEVESAVIRSPHFLDILSEIEBFTYTRAG 360

QY 361 LPLNTEYLEYVWGHSHIKYKNTNASSALERNYGTITSNKKYDYLANKDIPOVRSIGADL 420
DB 361 LPLNTEYLEYVWGHSHIKYKNTNASSALERNYGTITSNKKYDYLANKDIPOVRSIGADL 420

QY 421 ANYAQVGVFPVASTLLDKNTGSGSVGGFTYSKPHMTTMOVCTQNYNTIDEIPPENEPLS 480
DB 421 ANYAQVGVFPVASTLLDKNTGSGSVGGFTYSKPHMTTMOVCTQNYNTIDEIPPENEPLS 480

QY 481 RGYSHRLSHITSYSPSKNASSPARVGNLVPFAWTHRSADVNTVYSDKIQIPVYKAHTL 540
DB 481 RGYSHRLSHITSYSPSKNASSPARVGNLVPFAWTHRSADVNTVYSDKIQIPVYKAHTL 540

QY 541 VSGTIVIKGPGTGNILKRTSSGPLAYTSVSVKPLSORVBARIRYASTTNLRLPVITIS 600
DB 541 VSGTIVIKGPGTGNILKRTSSGPLAYTSVSVKPLSORVBARIRYASTTNLRLPVITIS 600

QY 601 GTRIYSINVKTNKMGDDLTFTNTFLATIGTAFPTSNYSDSLTVGADSPASGGEVYVDKF 660
DB 601 GTRIYSINVKTNKMGDDLTFTNTFLATIGTAFPTSNYSDSLTVGADSPASGGEVYVDKF 660

QY 661 ELIPVNPATPEAEEDLVAKKANVGLFTSKKDALQTSVTDYQVQANLVCECLSDLYPNE 720
DB 661 ELIPVNPATPEAEEDLVAKKANVGLFTSKKDALQTSVTDYQVQANLVCECLSDLYPNE 720

QY 721 KEMLDWAKKAEKLVQARNLLODTGFNRINGENGTGSGIEVAGDVLFPKDRSLRLTSA 780
DB 721 KEMLDWAKKAEKLVQARNLLODTGFNRINGENGTGSGIEVAGDVLFPKDRSLRLTSA 780

QY 781 REIDTETPTLYLQIQIDESLLKPYTRYKLGFGSSQDLIEIKLIHRANQIVKNVFDNLL 840
DB 781 REIDTETPTLYLQIQIDESLLKPYTRYKLGFGSSQDLIEIKLIHRANQIVKNVFDNLL 840

QY 841 PDVLVNSCGGIDRCSEQQYVDANTLALENNNGENGMSSDHPHIDTGEIDLNENGTI 900
DB 841 PDVLVNSCGGIDRCSEQQYVDANTLALENNNGENGMSSDHPHIDTGEIDLNENGTI 900

QY 901 WVVPKIPPTNGYATLGNLEVEEGLPSGETTLERAOQOECQWQDWARKEGASEKAYIAAK 960
DB 901 WVVPKIPPTNGYATLGNLEVEEGLPSGETTLERAOQOECQWQDWARKEGASEKAYIAAK 960

QY 961 QAIDRLFADYQOQKLSNGVEMSLAAQNLVOSIPVYVNDALPEIPGMNYTSTFELTNRL 1020
DB 961 QAIDRLFADYQOQKLSNGVEMSLAAQNLVOSIPVYVNDALPEIPGMNYTSTFELTNRL 1020

QY 1021 QQAMNLYDLERNAI PNGDFRNLGSLDNWATS DVNQQLSDTSLVLPVNNNSQVSOQFTVQPN 1080
DB 1021 QQAMNLYDLERNAI PNGDFRNLGSLDNWATS DVNQQLSDTSLVLPVNNNSQVSOQFTVQPN 1080

QY 1081 YRYVLRTVTRARKEGVDGVYIIRDGANQTEITLTFNICDDDTGVLSDADQTSYITKTVETPS 1140
DB 1081 YRYVLRTVTRARKEGVDGVYIIRDGANQTEITLTFNICDDDTGVLSDADQTSYITKTVETPS 1140

QY 1141 TEQWIDMSSETGVFNIESVELVLEE 1167
DB 1141 TEQWIDMSSETGVFNIESVELVLEE 1167

RESULT 2
AAR28900
ID AAR28900 standard; protein; 1157 AA.
XX
AC AAR28900;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-APR-1993 (first entry)
XX
DE Toxin 50C.
XX
KW Endotoxin; acarides; pest; Two Spotted Spider; mite; phytophagus.
XX
OS Bacillus thuringiensis; kumamotoensis PS50C.
XX
PN W09219106-A1.
XX
PD 12-NOV-1992.
XX
PF 30-APR-1992; 92WO-US003546.
XX
PR 30-APR-1991; 91US-00693210.
PR 13-SEP-1991; 91US-00759248.
PR 30-SEP-1991; 91US-00768141.
XX
PA (MYCO ) MYCOGEN CORP.
XX
PI Payne JM, Cannon RJC, Bagley AL;
XX
WPI; 1992-398411/48.
DR N-PSDB; AAQ30821.
XX
PT New Bacillus thuringiensis isolates and toxins - used for controlling acarid pests of livestock, fowl, stored prods. and plants.
XX
PS Claim 16; Page 53 + 49-51; 62pp; English.
XX
CC Gene sequences encoding a toxin which is active against acarides and is obtainable from B. thuringiensis isolates PS17a, PS17b, 33f2, PS2A1, PS69D1, PS6A1 and PS50C are given in AAQ30803-07 and AAQ30820-21 respectively. The toxin is a delta-endotoxin active against acarid pests, including the Two Spotted Spider mite. The isolates can be used against non-phytophagus mites such as acarid pests of livestock, fowl and stored prods. The genes can be cloned and used to transform other hosts, which can be used to control mites, or in the case of transgenic plants, be resistant to mites. (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 1157 AA;

Query Match      66.3%; Score 4008; DB 2; Length 1157;
Best Local Similarity 66.1%; Pred. No. 1.7e-255;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEVEIILDASSSTSVSDNSVRYPLANDQTTTLQNNMYKDYLRMSGEGNPELFGNP 60
DB 1 MSPNNQNEVEIILDASSSTSVSDNSVRYPLANDQTTTLQNNMYKDYLRMSGEGNPELFGNP 60

QY 61 ETPFISSTVQTGIGIVGVLGALGVPFAGQIASFYSTVGQWPSTSVSWEMIMKQVED 120
DB 61 ETPFISSTVQTGIGIVGVLGALGVPFAGQIASFYSTVGQWPSTSVSWEMIMKQVED 120

QY 121 LIDQKITDSVRKTAAGLGLGDLVDVYQKSLKNWLENRNDTRARSVVVTVYIALLELDFV 180
DB 121 LIDQKITDSVRKTAAGLGLGDLVDVYQKSLKNWLENRNDTRARSVVVTVYIALLELDFV 180

QY 181 AKIPSPAISSQGVPLLSVYQAQANLHLILRRDASIPGAEWGFTGCEISTFYDRQVTRTAQ 240
DB 181 AKIPSPAISSQGVPLLSVYQAQANLHLILRRDASIPGAEWGFTGCEISTFYDRQVTRTAQ 240

QY 241 YSDYCVKMYNTGLDKLKGNTNAASWLKYHQFREMTLLVLDLVALPNDYDTRTYPIETTAQ 300
DB 241 YSDYCVKMYNTGLDKLKGNTNAASWLKYHQFREMTLLVLDLVALPNDYDTRTYPIETTAQ 300

QY 301 LTRVYVTDPIVNRRTSGGFCRRWSLNSDISPSEVESAVIRSPHFLDILSEIEBFTYTRAG 360
DB 301 LTRVYVTDPIVNRRTSGGFCRRWSLNSDISPSEVESAVIRSPHFLDILSEIEBFTYTRAG 360

QY 361 LPLNTEYLEYVWGHSHIKYKNTNASSALERNYGTITSNKKYDYLANKDIPOVRSIGADL 420
DB 361 LPLNTEYLEYVWGHSHIKYKNTNASSALERNYGTITSNKKYDYLANKDIPOVRSIGADL 420

QY 421 ANYAQVGVFPVASTLLDKNTGSGSVGGFTYSKPHMTTMOVCTQNYNTIDEIPPENEPLS 480
DB 421 ANYAQVGVFPVASTLLDKNTGSGSVGGFTYSKPHMTTMOVCTQNYNTIDEIPPENEPLS 480

QY 481 RGYSHRLSHITSYSPSKNASSPARVGNLVPFAWTHRSADVNTVYSDKIQIPVYKAHTL 540
DB 481 RGYSHRLSHITSYSPSKNASSPARVGNLVPFAWTHRSADVNTVYSDKIQIPVYKAHTL 540

QY 541 VSGTIVIKGPGTGNILKRTSSGPLAYTSVSVKPLSORVBARIRYASTTNLRLPVITIS 600
DB 541 VSGTIVIKGPGTGNILKRTSSGPLAYTSVSVKPLSORVBARIRYASTTNLRLPVITIS 600

QY 601 GTRIYSINVKTNKMGDDLTFTNTFLATIGTAFPTSNYSDSLTVGADSPASGGEVYVDKF 660
DB 601 GTRIYSINVKTNKMGDDLTFTNTFLATIGTAFPTSNYSDSLTVGADSPASGGEVYVDKF 660

QY 661 ELIPVNPATPEAEEDLVAKKANVGLFTSKKDALQTSVTDYQVQANLVCECLSDLYPNE 720
DB 661 ELIPVNPATPEAEEDLVAKKANVGLFTSKKDALQTSVTDYQVQANLVCECLSDLYPNE 720

QY 721 KEMLDWAKKAEKLVQARNLLODTGFNRINGENGTGSGIEVAGDVLFPKDRSLRLTSA 780
DB 721 KEMLDWAKKAEKLVQARNLLODTGFNRINGENGTGSGIEVAGDVLFPKDRSLRLTSA 780

QY 781 REIDTETPTLYLQIQIDESLLKPYTRYKLGFGSSQDLIEIKLIHRANQIVKNVFDNLL 840
DB 781 REIDTETPTLYLQIQIDESLLKPYTRYKLGFGSSQDLIEIKLIHRANQIVKNVFDNLL 840

QY 841 PDVLVNSCGGIDRCSEQQYVDANTLALENNNGENGMSSDHPHIDTGEIDLNENGTI 900
DB 841 PDVLVNSCGGIDRCSEQQYVDANTLALENNNGENGMSSDHPHIDTGEIDLNENGTI 900

QY 901 WVVPKIPPTNGYATLGNLEVEEGLPSGETTLERAOQOECQWQDWARKEGASEKAYIAAK 960
DB 901 WVVPKIPPTNGYATLGNLEVEEGLPSGETTLERAOQOECQWQDWARKEGASEKAYIAAK 960

QY 961 QAIDRLFADYQOQKLSNGVEMSLAAQNLVOSIPVYVNDALPEIPGMNYTSTFELTNRL 1020
DB 961 QAIDRLFADYQOQKLSNGVEMSLAAQNLVOSIPVYVNDALPEIPGMNYTSTFELTNRL 1020
```


Db 61 ETFISSSTQTGTGIVGRIGALGFASQVYFVQVLPWSPKSDVINGEINERVE 120
Qy 121 LIDQKITDSVRKTAGLGLDGLVQYKSLKNWLENENDTRARVVVVVQYIALDLDFV 180
Db 121 LVDQKIEKYVOKALAEKGLGNALDVOYQSLDLENENDTRARVVVVVQYIALDLDFV 180
Qy 181 AKIPSPAISQVQVPLSVYAOANLHLLLRDASIFGAEWGFTPGGISTFYDQVTRTAQ 240
Db 181 SSIPSPAVSGHEVLLAVYAQAVNLHLLLRDASIFGEWGFPTGGEISRYNROVQLTAE 240
Qy 241 YSDYCVKNTGLDKGNALASWLYKHQFRREMTLLVLDLVALFNYDTRTPPIETTAQ 300
Db 241 YSDYCVKNTGLDKGNALASWLYKHQFRREMTLLVLDLVALFNYDTRTPPIETTAQ 300
Qy 301 LTRVVYTDPIVFNRETSQVFCRWSLNSDISPSEVESAVIRSPHLPDILSEIRFYTRAG 360
Db 301 LTRDVTTDPIAVNIVTSGFCPWSHSGILFTEVENNVIRPHLPDILSSVEINTSGG 360
Qy 361 LPLNTEYLEYVWGHISIKYKNT-NASSALERNYGTITSNKIKYDILANKDIPOVRSLGAD 419
Db 361 ITLANDAYINYSGHTLYRRRTADSTVTVTANYGRITSEK-NSPALERDRIPEINSTVAN 419
Qy 420 LANYTAQVGVYASFTLIDKNTGSGVGGFTYKPHHTMQVCTQNYNTIDEIPPERE-P 478
Db 420 LANYTQKAYGVPGSWFMYKR--GTSSTTAYLYSKTHLTALQCTQVYSSDEIPLDRTVP 477
Qy 479 LSRGVSHRLSHITSYSPKASSPARYGNLPVPAWTHRSADVNTNTVYSDIKITQIPVVKAH 538
Db 478 VASYSHRLSHITSYSPKASSPARYGNLPVPAWTHRSADVNTNTVYSDIKITQIPVVKAH 538
Qy 539 TLVSGTIVIKGPGFTGNILKRTSSGFLAYTSVSKSPISQRYRARIYASTNLTALFTY 598
Db 536 MLYLGSSVVGPGFTGDDILKRTNPSILGTFVAVTVNGSLSQRYRARIYASTNLTALFTY 594
Qy 599 ISGTRIYSINVKNTMKGDDLTNTFDLATIGTAPTPSNYSLSLTVGADSPASGGEVYVD 658
Db 595 YLGDITIEKRNFRKNDGASLTAYETPKFASFTDFQFRETQDKILLSMGDFSSGQEVYD 654
Qy 659 KPELIPVNAATPEARDLDAKAVNGLFTSKDALQTSVTDYOVNOAANLVCLSDELVP 718
Db 655 RLIEFIPVDITYEAEQDLAKAVNALFTNTKDLRPGVTDTEVNOAANLVCLSDDLTP 714
Qy 719 NEKRLMDAVKAKRLVQARNLLQDTGFNRIENGWGTSGTIEVAGDVLFDKRSRLIT 778
Db 715 NEKRLMDAVKAKRLVQARNLLQDTGFNRIENGWGTSGTIEVAGDVLFDKRSRLIT 774
Qy 779 SARBITETPTLYQIDDESLLKPYTRYKLGPIGSSQDLLEIKLIRHANQIVKNVPON 838
Db 775 GARBITETPTLYQIDDESLLKPYTRYKLGPIGSSQDLLEIKLIRHANQIVKNVPON 834
Qy 839 LLPDVLVNSCGIDRCSSQVVDANLALNNGENGNMSSDSHAFSDTGEIDLNENT 898
Db 835 LLPDVLVNSCGIDRCSSQVVDANLALNNGENGNMSSDSHAFSDTGEIDLNENT 888
Qy 899 GTWVFKIPTINGYATLGNLELVEBGLSGETLERAQOQOQWQKARKGASKAYYA 958
Db 889 GTWVFKIPTINGYATLGNLELVEBGLSGETLERAQOQOQWQKARKGASKAYYA 948
Qy 959 AKQALDRPADYQDQKLNKGVMSDMLAAQNLVQSIPIVYNDALPEIPGANTYSPTELTN 1018
Db 949 SKQAVRLYADYQDQKLNKGVMSDMLAAQNLVQSIPIVYNDALPEIPGANTYSPTELTN 1008
Qy 1019 RLOQANWLYDLNAPNGDFRNLGNLNNATPGVEVQOQINHTSVLVPNNWQVSOQFTVO 1078
Db 1009 RLOQANWLYDLNAPNGDFRNLGNLNNATPGVEVQOQINHTSVLVPNNWQVSOQFTVO 1068
Qy 1079 PNRYVLRVTAKEGVGQGVYIRDCANOTETLTTFMCDTGVLSADQTSYITKTVET 1138
Db 1069 PNRYVLRVTAKEGVGQGVYIRDCANOTETLTTFMCDTGVLSADQTSYITKTVET 1128
Qy 1139 PSTEQWIDMSTEGVFNTIESVELVLEE 1167
Db 1139 PSTEQWIDMSTEGVFNTIESVELVLEE 1167

Db 1129 PYTDOMMIEISETEGTFTYIESVELIVDVE 1157
RESULT 3
AAR25997
ID AAR25997 standard; protein; 1157 AA.
XX
XX AAR25997;
XX AC AC
XX 25-MAR-2003 (revised)
XX 27-JAN-1993 (first entry)
XX Delta-endotoxin.
XX B.t.PS50C; colorado potato beetle; Leptinotarsa decemlineata;
XX Pseudomonas; expression vector.
XX Bacillus thuringiensis.
XX EP498537-A2.
XX 12-AUG-1992.
XX 16-JAN-1992; 92EP-00300366.
XX 16-JAN-1991; 91US-00642112.
XX (MYCO) MYCOGEN CORP.
XX Poncerrada L, Sick AJ, Payne JM;
XX WPI; 1992-270497/33.
XX P-PSDB; AAR25997.
XX New insecticidal Bacillus thuringiensis strain NRRL B-18746 - and toxin,
XX Coleoptera.
XX Claim 6; Page 11-14; 16pp; English.
XX The sequence given is an approx. 130 kD protien known as Bacillus
XX thuringiensis PS50C (B.t.PS50C). This novel delta-endotoxin has been
XX shown to be active against the Colorado potato beetle (leptinotarsa
XX decemlineata). The gene encoding this toxin can be transformed into
XX suitable hosts such as Pseudomonas, which can then be applied to the
XX environment of coleopteran insects where they will proliferate and be
XX ingested by the insects. This DNA sequence can be inserted into an
XX expression vector which contains a promoter/operator region, a ribosome
XX binding site, polyadenylation signals, etc. This will allow transcription
XX and translation of the cells in the appropriate host. (Updated on 25-MAR-
XX 2003 to correct PN field.)
XX Sequence 1157 AA;
Query Match 66.3%; Score 4007; DB 2; Length 1157;
Best Local Similarity 66.1%; Pred. NO. 1.9e-255;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;
Qy 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGRENPELPGNP 60
Db 1 MSPNNQNEVEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSGRENPELPGNP 60
Qy 61 ETFISSSTQTGTGIVGVQVGLGALGFASQVYFVQVLPWSPKSDVINGEINERVE 120
Db 61 ETFISSSTQTGTGIVGVQVGLGALGFASQVYFVQVLPWSPKSDVINGEINERVE 120
Qy 121 LIDQKITDSVRKTAGLGLDGLVQYKSLKNWLENENDTRARVVVVVQYIALDLDFV 180
Db 121 LVDQKIEKYVOKALAEKGLGNALDVOYQSLDLENENDTRARVVVVVQYIALDLDFV 180
Qy 181 AKIPSPAISQVQVPLSVYAOANLHLLLRDASIFGAEWGFTPGGISTFYDQVTRTAQ 240
Db 181 SSIPSPAVSGHEVLLAVYAQAVNLHLLLRDASIFGEWGFPTGGEISRYNROVQLTAE 240

```
QY 241 YSDYCVKMYNTGLDCKGNTAASMLKYHQFRRMTLLVLDLVALPNNYDTRTYPLETTAAQ 300
DB 241 YSDYCVKMYKGLDCKAGTTSKMLNYHQFRRMTLLVLDLVALPNNYDTHYPLETTAAQ 300
QY 301 LTRVYTDPIVFNRETSGGFCRRMSLNSDISFSEVESAVIRSPHLDLILSEIEFYTTAAQ 360
DB 301 LTRVYTDPIAFNIVTSGFCNPNWTHSGILFYEVENNVRPPLHFDLILSSVEINTSRGG 360
QY 361 LPLNTEYLEYVWGHSHIKYKNT-NASSALERNYGTITSNKIKYYDLANKDIFQVRSIGAD 419
DB 361 ITLNDAYINWSGHTLKYRTADSTVITYTANYGRITSEK-NSPALEDRLDIFEINSTVAN 419
QY 420 LANYAQQVGVPPYASFTLLDKNTGSGVGGFTYSKPHHTMQVCTQNYNTIDIEIPPENE-P 478
DB 420 LANYQKAYGVGSPFWHVR--GTSSITAYILSKTHALQCTQVYESSDEIPLDRTVP 477
QY 479 LSRGSHRLSHITSYFSRKNASSPARYGNLPVANTHRASDVNTVYYSKDKITQIPVVAH 538
DB 478 VAESYSHRLSHITSHSPKNGS--AAYGSPFPVFWHTSADLNNITYSKDKITQIPAVKGD 535
QY 539 TLVSGTVIKGPGTGGMLKRTSGPLAYTSVSVKSPLSQRYRIRYASTTNLRPLVT 598
DB 536 MLTYGGSVVQGGPGTGGDLKRTNPSILGTFAVTVNGSLQRYRIRYASTTDFE-FTL 594
QY 599 ISGTRIYSINVNKTKNGDDLTFTATIGTATFTFSNYSLSLVGADSPASGGGEVYVD 658
DB 595 YLGDITIEKNRPNKTDNGASLTETFKYPASFTDQFRETQDKILLSGDFSSGGQEVYID 654
QY 659 KPELLIPVNAFEABEDLVAKAVNGLFTSKDALQTSVTDYVQVQAANLVECLSDLYP 718
DB 655 RIEFIPVDETVEAQDLAAKAVNALFTNTDGLRPGVTDYEVNQAANLVECLSDDLYP 714
QY 719 NEKRLMDAVKARKLVQARNLQDQFNRIENGNGWTGCTGIEVABGDVLPKDRSLAT 778
DB 715 NEKRLFDVAREKRLSGARNLQDDQFQINGENGWAASTGIEIVEGDVAFKGYLRUP 774
QY 779 SAREIDTETPTLYQQIDESLLKPYTRYKLAGFTGSSQDLBKILIRHANOIVKNVDPN 838
DB 775 GAREIDTETPTLYQKVEEGLKPYTRYRLGFGVSSQGLEIYITRQTNRIKVNVEDD 834
QY 839 LLPDVLVNSCGIDRCSEQQVVDANLALNNGNGWSSDSHAFSPHIDGIEDLNENT 898
DB 835 LLPDVSVPNSDGSINRCSEQKYVNSRL-----EGENRSGDAHFSLPIDIGELDYNENA 888
QY 899 GTWVFKLPTTNGYATLGNLELVEGPLSGETLERAQQOQOQWQKARKGASAKYVA 958
DB 889 GLWUGFKLTPDGYATLGNLELVEGPLSGDALERLQREEQWKIQMTRREEDORRYWA 948
QY 959 AKQAIIDLPAFYQDQKLSGVMSDMLAAQNLVQSIPIVYNDALPEIPEGMNTSTFELTN 1018
DB 949 SKQAVDLVADYQDQNLNPDVEITDLTRAQDLIQSIPVYVNEKPEIPEGMNTKTELTLD 1008
QY 1019 RLOQAWNLVLENAIPNGDFRGLSDWATSQVNTQOOLSDFTSVLYIPIWNSQVSOQPTVQ 1078
DB 1009 RLOQAWNLVQDNALPNGDFRGLSNWATPGVEQQINHTSVLYIPIWNEQVSOQPTVQ 1068
QY 1079 PNRYVLRVTKRKGVDGYIIRDCANOTELTFNIDCDDTGVLSADQTSYITKVEFT 1138
DB 1069 PNQVLRVTKRKGVGNGYISIRGGNQSELTFTFSADYDTNGYNDQTYITKVTFI 1128
QY 1139 PSTEQWIDMSREGVFNIESVELVLEE 1167
DB 1129 PYTDQMWIEISREGTFVIESVELIVDVE 1157
RESULT 4
ID AAR33768
XX AAR33768 standard; protein; 1157 AA.
AC AAR33768;
XX DT 25-MAR-2003 (revised)
```

```
DT 15-JUL-1993 (first entry)
XX Bt isolate Ps50C.
KW Lepidopteran-active; toxin; Ti; Ri; plant; cell; Bc.
XX Bacillus thuringiensis.
OS WO9304587-A1.
PN 18-MAR-1993.
PD 11-SEP-1992; 92WO-US007697.
XX 12-SEP-1991; 91US-00758020.
PR 08-SEP-1992; 92US-00941650.
XX (MYCO ) MYCOGEN CORP.
XX Uyeda KA, Bradfish GA;
XX WPI; 1993-100566/12.
XX N-PSDB; AAQ38653.
XX Controlling lepidopteran pests - using compen. of Bacillus thuringiensis
XX strains or plants or microorganisms transformed with their toxin genes.
XX Claim 12; Page 22-25; 38pp; English.
XX The sequences given in AAR33768-70 represent lepidopteran-active toxins.
XX The DNA encoding these sequences were used within a Ti or Ri plasmid, to
XX transform plant cells. Whole plants can then be regenerated from the
XX transformed cells. The toxin may also be produced by cloning Bacillus
XX thuringiensis (Bt). It may then be applied directly to the plant locus.
XX (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to
XX correct PR field.)
XX SQ Sequence 1157 AA;
Query Match 66.3%; Score 4007; DB 2; Length 1157;
Best Local Similarity 66.1%; Pred. No. 1.9e-255;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;
QY 1 MSPNNQNYEILDASSSTSVSDNSVRYPLANDQTTTLQNNYKYDYLRMSEGENPELFGNP 60
DB 1 MSPNNQNYEIIIDATPSTSVSDSNRYFPANBPTDALQNNYKYDYLKMSGGENPELFGNP 60
QY 61 ETPFSSSTVQTCIGIVGVILGALGYPPAGQIASPYSPVIGQLWPSSTVSVWEMKQVED 120
DB 61 ETPFSSSTVQTCIGIVGVIRILGALGYPPAGQIASPYSPVIGQLWPSKSVVDINGEIMERVEE 120
QY 121 LIDQKITDSVRTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVQYIALDLFV 180
DB 121 LVDQKIEKYKDKALAEKGLGNALDGVYQOSLEDWLENENDARTSVSVNQFIADLNFV 180
QY 181 AKIPSPALSGQEVPLLSVYAQAANLHLILLRDASIFGAEWGTTPGISTFYDQVTRTAQ 240
DB 181 SSIPSPASVSGHEVLLAVYAQAVNLHLILLRDASIFGAEWGTTPGISTFYRQVLTAE 240
QY 241 YSDYCVKMYNTGLDCKGNTAASMLKYHQFRRMTLLVLDLVALPNNYDTRTYPIETTAQ 300
DB 241 YSDYCVKMYKGLDCKAGTTSKMLNYHQFRRMTLLVLDLVALPNNYDTHYPIETTAQ 300
QY 301 LTRVYTDPIVFNRETSGGFCRRMSLNSDISFSEVESAVIRSPHLDLILSEIEFYTTAAQ 360
DB 301 LTRVYTDPIAFNIVTSGFCNPNWTHSGILFYEVENNVRPPLHFDLILSSVEINTSRGG 360
QY 361 LPLNTEYLEYVWGHSHIKYKNT-NASSALERNYGTITSNKIKYYDLANKDIFQVRSIGAD 419
DB 361 ITLNDAYINWSGHTLKYRTADSTVITYTANYGRITSEK-NSPALEDRLDIFEINSTVAN 419
QY 420 LANYAQQVGVPPYASFTLLDKNTGSGVGGFTYSKPHHTMQVCTQNYNTIDIEIPPENE-P 478
DB 420 LANYQKAYGVGSPFWHVR--GTSSITAYILSKTHALQCTQVYESSDEIPLDRTVP 477
```


Db 949 SKQAVRLVADYQDLQLPNDVEITDLTAQDLIQSIPYVYNEKPEIPGMNTKFTELTD 1008
 Qy 1019 RLQQAANLYDLARNALPNCDFRGLSDWATSDVNVQQLSDTSVLVFNWNSQVSQOFTVQ 1078
 Db 1009 RLQQAANLYDQRNALPNCDFRGLSNWATPGVEVQQINHTSVLVIPNWDQVSQOFTVQ 1068
 Qy 1079 PNRYVLVLTARKEGVGGYVIRGANGOTETLTENICDDDTGVLSDQTSYITKTVFT 1138
 Db 1069 PNQRYVLVLTARKEGVGGYVIRGANGOTETLTENICDDDTGVLSDQTSYITKTVFT 1128
 Qy 1139 PSTQVWTDMSSETGVFNIESVELVLEE 1167
 Db 1129 PYTDQMIEISSETGTFYIESVELVLEE 1157

RESULT 8

AAR32354
 ID AAR32354 standard; protein; 1149 AA.

AC AAR32354;

XX 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 14-JUN-1993 (first entry)

XX

DE Coleoptera toxin from B.t. serovar japonensis variety Buihui.

XX Delta-endotoxin; larvae; Anomala cuprea; spore; crystal; Bacillus;

KW thuringiensis; virus; insect; FERM BP-3465.

XX

OS Bacillus thuringiensis; serovar japonensis variety Buihui.

XX

FN W09303154-A1.

XX

PD 18-FEB-1993.

XX

PF 31-JUL-1992; 92WO-US006404.

XX

PR 02-AUG-1991; 91JP-00193810.

XX

PR 23-JUL-1992; 92US-00915203.

XX

PA (MYCO) MYCOGEN CORP.

XX

PA (KUBI) KUBOTA CORP.

XX

PI Ohba M, Iwahana H, Sato R, Suzuki N, Ogiwara K, Sakanaka K;

XX

PI Hori H, Asano S, Kawasaki T;

XX

DR WPI; 1993-076511/09.

XX

DR N-PSDB; AAQ36866.

XX

PT New strain of Bacillus thuringiensis serovar japonensis - producing toxin

XX

PT active against coleoptera larvae.

XX

PS Claim 5; Page 28; 48pp; English.

XX

CC The protein sequence is that of a toxin active against Coleoptera that is produced from a pure culture of Bacillus thuringiensis serovar japonensis variety Buihui (FERM BP-3465). The toxin is a delta-endotoxin which has a mol. wt. of ca. 130 kD. It is useful for control of coleoptera larvae e.g. it is effective against Anomala cuprea but has little effect on Lepidoptera. The toxin can be used as B.t. spores or crystals, as opt. treated cells (B.t. or transformed microorganisms) or it is expressed by plants. (Updated on 25-MAR-2003 to correct FN field.) (Updated on 24-OCT-2003 to standardise OS field)

XX

SQ Sequence 1149 AA;

XX

Query Match 60.6%; Score 3662.5; DB 2; Length 1149;

XX

Best Local Similarity 63.8%; Pred. No. 1.1e-232; Indels 61; Gaps 21;

XX

Matches 752; Conservative 109; Mismatches 257;

XX

1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKYLRMSEGENPELFGNP 60

Db 1 MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKYLRMSEGENPELFGNP 60
 Qy 61 EFTISS-STVGTGIVGVQALCALGVPAGOLASPVYVIGVQWSPSSSTVSVVEMIMKQVE 119
 Db 61 GTFISAQAVAGTGDIVSTIISGIGIPVIGVFSILGSLIGLLWFSNNENWQIPANRVE 120
 Qy 120 DLIDQKIDTSVRKALAGLQGLDGLDLYVQSKLKNWLENRNDTRARSVVVYQYIALLEDF 179
 Db 121 ELIDQKILDSVRSRAIADLANSRJAVEYYQNALERDKRNPHSTRSAALVKERFCNAEAL 180
 Qy 180 VAKIPSPAI SQGVPLLSVYAAQANHLHLLLDASIFGAEWGFTPGELSTFYDQVTRTA 239
 Db 181 RTNMGSPSQNTYETPELLPYAQAASLHLLVRDVOIYCKGNGYPONDIDLFLYKBOVSYA 240
 Qy 240 QYSDYCVKWNVTGLDKLGTNAASVLYKHOFREMTLLVLDLVALFPNYDTYPIETTA 299
 Db 241 RYSDHCVOYINAGNLKLGAKQWYINRFRRENWVLDLVALFPNYDARIYPLETNA 300
 Qy 300 QLTREVTDPPIVFNRETSGGFCRRWSLN-----SDI-----SPSEVSAVIRSPHLFDIL 349
 Db 301 ELTREIPTDPV-----GSVVTGQSTLISWYDMIPAALPSFSTLEN-LLRKPOFFILL 352
 Qy 350 SEIBFTTRAGLPLNNT-EYLEVNVGHSIKYKNTNASSALEBNYGTITSNKIKYDLANK 408
 Db 353 QEIRMYTS---FRQNGTIEYNYMGQRLLTSIYIGSSF--NKYSGVLGAEDIIIPVQGN 407
 Qy 409 DIFQVRSIGADLANVYAGVYVYASFTLLDKNTGSGVGGFTYKSPHTTMOVCTQNYNT 468
 Db 408 DIYRV-VMTYIGRYTNSLLGVNVPVF-YFNSNTQK-----TYSKP-----KQFAGGIKT 454
 Qy 469 ID---EIPNEPEPLSRGSHRLSHITSYSPSKNASSPARYGNLVPFAMTHRSADVTNTVY 525
 Db 455 IDSGBELTYEN---YQSYSHRVSYITSPEIKSTGTV--LGVPPIFGMTHSSASRNNPIY 509
 Qy 526 SDKITQIPVVAHTLVSGTTVIKGPQ-PTGHNILKRTSSGPLAY---TSVSVKSPLSQRY 581
 Db 510 ATKISQIPINKASRTSGGAVNWFQGLYNGGVPWKLSGSGSQVINLRVAITDAGK-ASQRY 568
 Qy 582 RARIRYASTTNLRLFTIIS-----GTRIYSINNVKTNKGGDITFNFDLATIGTA-F 633
 Db 569 RIRIRYASDRAGK--TISRSPENPATYSIAIYNTMTSTWASLTYSYFAESGPINL 626
 Qy 634 TFSNYSDSLTVGADSPASGGVYVDKFLIPVNAATFEAEEDLDVAKAVNGLFTSKDAL 693
 Db 627 GISGSSRTFDISITKEGAANLYIDRIBFIPVTLTFAEEDLDVAKAVNGLFTSKDAL 686
 Qy 694 QTSVTDYQVNOAANLVECLSDLELYPNEKMLWDAVKEAKLYOARNLLQDTGPNRNGEN 753
 Db 687 QTSVTDYQVNOAANLIECLSDLELYPNEKMLWDAVKEAKLYOARNLLQDTGPNRNGEN 746
 Qy 754 GWTSGTGLEVAEGDVLFDORSRLTSAREIDTETPTLYLQOIDSLLKPYTRYKLKGP 813
 Db 747 GWTSGTGLEVEGSDVLFDORSRLTSAREIDTETPTLYLQOIDSLLKPYTRYKLKGP 806
 Qy 814 GSSQDLLEKLIRHRANQIVKXNVPDNLPLFVNSCGGIDRCSEQQYVDANALNNGEN 873
 Db 807 GSSQDLLEKLIRHRANQIVKXNVPDNLPLFVNSCGGIDRCSEQQYVDANALNNGEN 866
 Qy 874 GNMSSDSHAFPHIDTGBIDLNENTGIWVVKIPITNGVATIGNLEVEEGLSGETLER 933
 Db 867 GNMSSDSHAFPHIDTGBIDLNENTGIWVVKIPITNGVATIGNLEVEEGLSGETLER 926
 Qy 934 AQOQEQWQDMARKRGASKEYAQAQAIIDRLPADYQDQKLSGVNSDMLAAQNLVQS 993
 Db 927 AQOQEQWQDMARKRGASKEYAQAQAIIDRLPADYQDQKLSGVNSDMLAAQNLVQS 986
 Qy 994 IPYVYNDALPEIPGMNTYFTLNLRLQQAANLYDLRNAIPNGDFRNGLSQWNAISDVNV 1053
 Db 987 IPYVYNDALPEIPGMNTYFTLNLRLQQAANLYDLRNAIPNGDFRNGLSQWNAISDVNV 1046
 Qy 1054 QQLSDTSVLVLPNNSQVSQOFTVQPNRYVLVLTARKEGVGGYVIRGANGOTETLTENI 1113

Db 1047 QQLSDTSVLVFPNNSQVSOQFTVPQPNRYRLVRLVARTARKEGVGDGYVIIRDGANQTETLTP 1106
 Qy 1114 NICDDDTGVLSDADQTSYITKTVEFTPTSTEQVWIDMSETE 1152
 Db 1107 NICDDDTGVLSDADQTSYITKTVEFTPTSTEQVWIDMSETE 1145

RESULT 9
 AAR51692
 ID AAR51692 standard; protein; 1149 AA.
 XX AAR51692;
 AC AAR51692;
 XX 16-OCT-2003 (revised)
 DT 04-NOV-1994 (first entry)
 DE B.thuringiensis serovar Japonensis insecticidal protein.
 XX insecticidal protein; Coleoptera larvae; Buibui fungus.
 XX Bacillus thuringiensis; (serovar Japonensis).
 OS strain Buibui).
 XX Key Location/Qualifiers
 FT 751..766
 FT /note= "amino acids 751-766 are not given in the
 FT specification and so have been decoded from AAQ58975"
 XX JP06065292-A.
 PN 08-MAR-1994.
 XX 11-AUG-1992; 92JP-00213886.
 XX 11-AUG-1992; 92JP-00213886.
 XX (KUBI) KUBOTA CORP.
 PA WPI: 1994-121220/15.
 DR N-PSDB; AAQ58975.
 XX Insecticidal protein and DNA from Bacillus thuringiensis serovar
 PT Japonensis strain Buibui - useful in insecticides against Coleoptera
 PT insects.
 XX Claim 1; Page 9-13; 18pp; Japanese.
 XX This insecticidal protein has activity against Coleopteran insect larvae
 CC and has been isolated from Bacillus thuringiensis serovar japonensis
 CC strain Buibui. (Updated on 16-OCT-2003 to standardise OS field)
 XX Sequence 1149 AA;
 SQ

Query Match 60.6%; Score 3662.5; DB 2; Length 1149;
 Best Local Similarity 63.8%; Pred. No. 1.1e-232;
 Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

Qy 1 MSPNNQVEYILDASSTVSQVPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60
 Db 1 MSPNNQVEYILDASSTVSQVPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60
 Qy 61 ETPFIS-STVQTGIGVGVGLGVPFAGQIASFYFVGLWPSSTVSVMIMKQVE 119
 Db 61 GTPISAQDVGIDIVSTIGLGHVGLGVPFAGQIASFYFVGLWPSSTVSVMIMKQVE 120
 Qy 120 DLIDQKIDTSVRKALAGLQGLDGVYQKSLKNNLNNDTRARSVVVTVYALDLDF 179
 Db 121 ELIDQKIDTSVRKALAGLQGLDGVYQKSLKNNLNNDTRARSVVVTVYALDLDF 180
 Qy 180 VAKIPGAFSGQVPLLSVQAQANLHLLLRDASFGAEWFTPGESTFYDQVTRTA 239
 Db 181 RTNMGSGFSQNTVETPLLPYQAASLELLVVRDVQVYKGEWGPQNDIDLKYKEQVSITA 240

Qy 240 QYSDYCVKMYNTGLDKLGTNAASWLKYHQFPRREMTLLVLDLVALPQNYDTRTYPIETTA 299
 Db 241 RYSDHCQVQWYAGLKLGRGTGAQWVDYNRFRREMNVMVLDLVALPQNYDARIYPLETNA 300
 Qy 300 QLTREYVTPPIVFNRETSGPCRRWSLN-----SDI-----SSEVESAVIRSPHLDIL 349
 Db 301 ELTREIFTDPV-----GSYVTGQSTLISWYDMIPALPSPSTLEN-LLRKEDFPFTLL 352
 Qy 350 SEIEFTYTRAGLPLNNT-EYLEYVWVGHSIKYKNTNASSALERNYGTITTSNKIKYDILANK 408
 Db 353 OEIRMYTS---FRQNGTIEYNTWGGQRLTSLIYIGSSP--NKYSGVLAGAEDIIPVQCN 407
 Qy 409 DIPQVRSGLADLANIYAQVYVPYASFTLLDKNTSGSGVGGPYFKPHTTHQVCTQNYNT 468
 Db 408 DIYRV--VMYTIIGRYTNSLLGVNVPVTP-FYSNNTQK-----TYSK-----KOPAGGIKT 454
 Qy 469 ID---EIPPENEPLSRGYSHLSHITSYSPSKNASSPARYGNLPVPAWTHRSADVTWVY 525
 Db 455 IDSGEELTYEN---YOSYSHRVSIYITSFEIKSTGGTV--LGVFIPGWTSSASRNFIY 509
 Qy 526 SDKITQIPVVKATHLVSGTIVIKPGP-FTGKNILKRTSSGLPAY---TSVSVKSPLSQRY 581
 Db 510 ATKISOIPINKASRTSGGAVWVNFQBGLYNGPVPWKLSSGSQVNLRLVATDAKG-ASQRY 568
 Qy 582 RARIRVASTNLRFLFTVIS-----GTRIYSINVNKTWKGGDDLTFTNFDLATIGTA-P 633
 Db 569 RIRIRVADRAGKF--TISRSPENPATYSASLAYTNMTSNASLTYSSTAYAESGPINL 626
 Qy 634 TFSNYSDSLTVGADSPASGGEVTVDFELIPVATPEAREEDLVAKKAVNGLFTSKKDAL 693
 Db 627 GISGSSRTFDISITKEAGAAANLYIDRIEFPVNTLFEAREEDLVAKKAVNGLFTNEKDAL 686
 Qy 694 QTSVTDYQVNOAANLVECLSDLEYFNEKRLMDAVKEAKELVQARNLLQDTGFNRINGEN 753
 Db 687 QTSVTDYQVNOAANLVECLSDLEYFNEKRLMDAVKEAKELVQARNLLQDTGFNRINGEN 746
 Qy 754 GMTGSTGIEVAGDVLFKDRLTSAREIDTETPTLYLQOIDEISLLKPYTRYKLKGF 813
 Db 747 GMTGSTGIEVAGDVLFKDRLTSAREIDTETPTLYLQOIDEISLLKPYTRYKLKGF 806
 Qy 814 GSSQDLLEIKLIRHRANOIVKQVNDLLPDVLPVNSCGGIDRCSEQVVDANLLENNGEN 873
 Db 807 GSSQDLLEIKLIRHRANOIVKQVNDLLPDVLPVNSCGGIDRCSEQVVDANLLENNGEN 866
 Qy 874 GNMSSDSHAFSFHIDTGEIDLANTGTIWWVFKIPTTNGVATLGNLELVBBGPLSGETLER 933
 Db 867 GNMSSDSHAFSFHIDTGEIDLANTGTIWWVFKIPTTNGVATLGNLELVBBGPLSGETLER 926
 Qy 934 AQOQEQWQDKMARKGASEKAYAAKQAIIDRLPADYQDOKLNSGVMSDLAAQNLVQS 993
 Db 927 AQOQEQWQDKMARKGASEKAYAAKQAIIDRLPADYQDOKLNSGVMSDLAAQNLVQS 986
 Qy 994 IPVYNDALPEIRGMNYTSFTELNRLLQAAWNLVDLNRNIPNGDFRNLGSDMNATSDVNV 1053
 Db 987 IPVYNDALPEIRGMNYTSFTELNRLLQAAWNLVDLNRNIPNGDFRNLGSDMNATSDVNV 1046
 Qy 1054 QQLSDTSVLVFPNNSQVSOQFTVPQPNRYRLVRLVARTARKEGVGDGYVIIRDGANQTETLTP 1113
 Db 1047 QQLSDTSVLVFPNNSQVSOQFTVPQPNRYRLVRLVARTARKEGVGDGYVIIRDGANQTETLTP 1106
 Qy 1114 NICDDDTGVLSDADQTSYITKTVEFTPTSTEQVWIDMSETE 1152
 Db 1107 NICDDDTGVLSDADQTSYITKTVEFTPTSTEQVWIDMSETE 1145

RESULT 10
 AAR59256
 ID AAR59256 standard; protein; 1210 AA.
 XX AAR59256;
 AC AAR59256;
 XX 07-OCT-2002 (first entry)
 XX

Bacillus thuringiensis Cry1219-2 protein sequence.

Pesticidal; spraying; dusting; broadcasing; seed coating; insect pest;
Colorado potato beetle; western corn rootworm; southern corn rootworm;
insect target range; endotoxin; Cry1218.

Bacillus thuringiensis..

WO200234774-A2.

02-MAY-2002.

24-OCT-2001; 2001MO-US045468.

24-OCT-2000; 2000US-0242838P.

23-OCT-2001; 2001US-00032717.

(DUPO) DU PONT DE NEMOURS & CO E I.

Abad AR, Duck NB, Feng X, Flannagan RD, Kahn TW, Sims LE;

WPI; 2002-519178/55.

N-PSDB; ABK87235.

New isolated pesticidal polypeptide useful for impacting insect pest e.g.
Colorado potato beetle.

Claim 4; Page 103-105; 176pp; English.

The present invention relates to a new pesticidal polypeptide. The invention is useful for impacting an insect pest by applying the molecules of the invention to the environment of the insect pest by spraying, dusting, broadcasing, or seed coating, where the insect pest is selected from Colorado potato beetle, western corn rootworm or southern corn rootworm. The invention is also useful for increasing insect target range and for producing transgenic microorganisms and plants that express the pesticidal polypeptide. The invention is also useful for producing transformed plants and in transforming any organism to produce the pesticidal polypeptide of the invention. The present amino acid sequence represents a *Bacillus thuringiensis* wild-type Cry1218 endotoxin protein

Sequence 1210 AA;

Query Match 54.0%; Score 3263.5; DB 5; Length 1210;

Best Local Similarity 53.4%; Pred. No. 2.6e-206;

Matches 667; Conservative 178; Mismatches 284; Indels 119; Gaps 21;

1 MSPNNQNEYELDASSSTVSNDNRYPLANDQTTTLQNNYKDYLRMSGENPELFGNP 60

1 MSPNNQNEYELDASSSTVSNDNRYPLANDQTTTLQNNYKDYLRMSGENPELFGNP 60

61 ETPIS-SSTVGTGIGVGVGLGVPAGQIASFVIGVQLPSSVTSVWEMIMKQVE 119

61 EVLVSGQDAKAAIDVGLKLSGLGVFPVGVLSVLTQLDILPFGSKQSWEIFMEQVE 120

120 DLIDQKITDSVRKTAAGLQGLDGLDYQKSLRNLENRNDTRARSVVVYQYIALSLDF 179

121 ELINQKIAEYARNKALSELEGNNYQLYLTALSEWKENPNSRDLVRNRPEILDLP 180

180 VAKIPSPAIQGVPLLSVVAQAANLHLLLRDASIFGAEWGFTPGELSTFYDQVTRTA 239

181 TQYMFSPRVNTFVPLFTVYTTQAANLHLLKADASIFGEWGWSTTTNNYDQMKLTA 240

240 QYSDCYKRVNTGDKLKGTAASWLYKHQPRREMTLLVLDVALFNYDTRTYPIETTA 299

241 EYDHCWKVYETGLAKLKGTSKAKQVNDVQPRREMTLLVLDVALFNYDTRTYPIETTA 300

300 QLTREVTYDPIVFNRETSGGFCRRWSLNSDISFSEVSASVIRSPHLEDFILSEIFYTTTA 359

301 QLTREVTYDPIGAVNVSSIG---SW-YDKAPSGVIESVIRPVRVFDYITGLVYITQSR 356

360 GLPLNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIKYDLANKDFQVRSIGA 418

357 S--ISSARYIRHWAGHOISYHRIPSDNIIKQMYGTQNLHSTSTPFDFTNYDIYKTLSDA 414
419 DLAN-----YAAQVYGVYASFTLLDK--NTGSSGVGGFTYSKPHTTMQVCTONTTID- 470
415 VLLDIVPFGTYIYFFGMPEVEFFMVQNLNTRK-----TLKTNPVSK 456
471 -----EIPPE--NEPLSRGYSHRLSHITSYFSKVNASSPARYGNLPVPAMTHRS 517
457 DIIAGTRDSLELPPETSQPNVESYSHRLCHITSIPATGSTT-----GLVPVFSWTHRS 511
518 ADVNTVYSDKITQIPVVKAHTL---VSG---TTVIGPGFTGNNILKRTSSGGLAYTSV 571
512 ADLINAHSKTIQIPVVKVSDLAPSITGGPNNTVWSGPGFTGGIIRKIVNG-VIIISHM 570
572 SVK-SPLSORYRARIYASTTNLRKFVTISIRYISINVKTNKNGDGLTNTFDLATIG 630
571 RVKISDINKYSMRIRYASANTFYPINPSENVKS-HAQTMRGALTYNKFNATL- 628
631 TAFTFSNYSDLTVQADSPAS---GGEVYVDKFLIPVNAITFEAEEDLDVAKKAVNGLPT 687
629 PPIKFTTTEPPIITLGAIFEAEDFLGIBAYIDRIEPIPVDETYEAEQDLAAKAVNALFT 688
688 SKDQALQTSYTDYQVNOAANLVECLSDLEYPNEKGLMDAVKEAKELVQARNLLQDTGPN 747
689 NTKDGLPFGVTDYEVNOAANLVECLSDLLYPNEKRLFLDAVREAKRLSEARNLLQDDPQ 748
748 RINGENGWGTSGTGLEVAEGDVLKDRSLRITSAREIDTETPTLYLYQDIDSLKLPYTRY 807
749 EINGENGWTAAGTSLGIEVTEGDLAFGRVLRPLPGAREIDTETPTLYLYQKVEGVLKPYTRY 808
808 KLRGFIGSSQDLKILRHRANQIKVNPVNDLPLVLPVNSCGGIDRCSEQQYVDANLAL 867
809 RLRGFVSSQGLAIFTIRHQTNRIVKNVPDOLLPDVSPVNSDGSINRCSEKQYVNSRLEV 868
868 ENNGENGWMSDASHAFSHIDTGHIDNENTGIWVFKIPTNGYATLGNLELVEEGPLS 927
869 ENR-----SGEAHEFSIPIDTGEIDYNENAGIWWGFKITDPEGYATLGNLELVEEGPLS 922
928 GETLRAQQOQOQWODKARKGASEKAYVAAKQAIIDRLPADYQDQKLSNGVMSDMLAA 987
923 GDALERLQREEQWKKIQNTRRBETDRRYWASKQAVDLVADYQDQQLNPDVEITDLTAA 982
988 QNLVQSIPYVYNDALPEIPGMYNTSFTELTNRLOQAANLVDLRNAIENGDFRNGLSDNA 1047
983 QDLIQSIPIVYVYVNEFPPEIPGMYNTTKFTELTDLQQAAMSLYDQRNAIPNGDFRNGLSDNA 1042
1048 TSDVNVQOOLSQTSVLVNPVNSQVSOQTPQPNRYVYLVRTARKEGVDGVVIRDDGANO 1107
1043 TPGVEVQOQIHTSVLVIPNMDQVSOQTPQPNRYVYLVRTARKEGVDGVVIRDDGANO 1102
1108 TETLTFTNICDDDT-GVLSA----- 1125
1103 TETLTFTSASDYDNGMYNTQVNTNGYNTNAYNTQASSTNGYNNANNMYNTQASNTNGYN 1162
1126 -----DQSYIYNTKFTSTEQWIDMSETEGVFNIESVELVLEE 1167
1163 TNSVYNDQTYITNTVTTFIPTYDQWIMSETEGTFTIESVELIVDVE 1210

RESULT 11

AAU99255

ID AAU99255 standard; protein; 1206 AA.

XX AAU99255;

XX 07-OCT-2002 (first entry)

XX *Bacillus thuringiensis* Cry1218-1 protein sequence.

XX Pesticidal; spraying; dusting; broadcasing; seed coating; insect pest;

XX Colorado potato beetle; western corn rootworm; southern corn rootworm;

XX insect target range; endotoxin; Cry1218.

XX	Bacillus thuringiensis.	116	DB	415	VLLDIYVPGYVIFPGMPVEFPMVQLNNTK-----TLKYNFVSK	456
OS			QY	471	-----EIPPS--NEPLSRGSHRLSHITSYSFSKNASSPARYMGLPVPANTHRS	517
XX	WO200234774-A2.		DB	457	DIIASTRDSELELPETSQPNVYESYHRLCHITSIPATGNTT-----GLVPFVSWTHRS	511
PD	02-MAY-2002.		QY	518	ADVTNIVSDKLTQTPVKAHTLVSGTIVIKGPGTGEMILKETSS-----GPLAYTSVS	573
XX	24-OCT-2001; 2001WO-US045468.		DB	512	ADLNTIYSDKLTQTPAVRCMDNLPFPVVKPGHGTGGDLLQYNRSTGSGVTLFLARYGL	571
XX	24-OCT-2000; 2000US-0242838P.		QY	574	KSPLSORYRARIYASTTNLRFPVTISGTRIVSNVKNKWKGGDLTFTFIDLATIGTAF	633
PR	23-OCT-2001; 2001US-00032717.		DB	572	ALEKAKYRVRURYATDADIVLHND-----QIOMPKNPFGEDJTSKTFKVADAITL	626
XX	(DUPO) DU PONT DE NEMOURS & CO E I.		QY	634	---TFNSYSDSLTVGADSPAS--GGEVYVDKFLIPVNATFEABEDLDVAKAVNGLPSTK	689
PI	Abad AR, Duck NB, Feng X, Flannagan RD, Kahn TW, Sims LE;		DB	627	NLATDSSLAKENLGEDPNSTLSGIVYVDRIEFIPVDETVEABQDLAAKAVNALPTMT	686
XX	WPI: 2002-519178/55.		QY	690	KDALQTSVTDYVQNAANLVECLSDLYPNKXEMLDVAKEALVQARNLLQDTPGNRI	749
DR	N-PSDB; ABR67234.		DB	687	KDGLRPGVTDYEVNQNAANLVECLSDLYPNKXRLFLDVAKEALVQARNLLQDTPQEI	746
XX	New isolated pesticidal polypeptide useful for impacting insect pest e.g. Colorado potato beetle.		QY	750	NGENGWTSGTGLEVAEGDVLFXDRSLRLTSAREIDTETPTLYLQQIDESLLAPYTRYKL	809
PS	Claim 4; Page 96-98; 176pp; English.		DB	747	NGENGWTSGTGLEVAEGDVLFXDRSLRLTSAREIDTETPTLYLQQIDESLLAPYTRYKL	806
CC	The present invention relates to a new pesticidal polypeptide. The invention is useful for impacting an insect pest by applying the molecules of the invention to the environment of the insect pest by spraying, dusting, broadcasting, or seed coating, where the insect pest is selected from Colorado potato beetle, western corn rootworm or southern corn rootworm. The invention is also useful for increasing insect target range and for producing transgenic microorganisms and plants that express the pesticidal polypeptide. The invention is also useful for producing transformed plants and in transforming any organism to produce the pesticidal polypeptide of the invention. The present amino acid sequence represents a Bacillus thuringiensis wild-type Cry1218 endotoxin protein		QY	810	KGFIGSSQDLKILNHRANQIVKNVDPNLLPDLVFNKSGGIDRCSEQQYVDANALEN	869
CC			DB	807	RGFVSSQGLEIFTRHQTNRIVKNVDPNLLPDLVFNKSGGIDRCSEQQYVDANALEN	866
CC			QY	870	NGENGWSSDSHAFPHIDTGDIDNENGIWVFKIPTNGVATLGNLSEVEGPLSGE	929
CC			DB	867	R-----SGEHEFSIPIDTGEIDYENAGIWWGPKITDPEGVATLGNLSEVEGPLSGD	920
CC			QY	930	TLERAOQOQOQWQDKMKRGKSGAKYAAKQAI DRLFADYQDKLNSGVEMSLAAQN	989
CC			DB	921	ALERLQREQQWQKIQTRREEDRYMASKQAVDRLYADYQDQLNPDVEITDLTAQD	980
CC			QY	990	LVSQIPYVYNDALEIPGNYTSFTLTLNRLQOAWNLXDLRNAIPNGDFNGLSDWNATS	1049
CC			DB	981	LQISPIVYNEMFPEIPGMYTKFTLTLNRLQOAWNLXDLRNAIPNGDFNGLSDWNATS	1040
CC			QY	1050	DYNVQQLSDTSVLVFNKNSQVSOQFTVQPNRYLRTARKEGCGYVIRDGANQTE	1109
CC			DB	1041	GVEVQOQINETSVLVFNKNSQVSOQFTVQPNRYLRTARKEGCGYVIRDGANQTE	1100
CC			QY	1110	TLTENICDDDT--GVLSA-----	1125
CC			DB	1101	TUTFSASDITDGTNGYNTQVNTNGYNTNAYNTQASSTNGYNNANNMYNTQASNTGNTN	1160
CC			QY	1126	----DQTSYITTKTVEFTPESTEQVWIDMSETEGVFNIESVELVLEE	1167
CC			DB	1161	SVYNDQGTITTKTVPFIPYTDQMWIEMSETEGVFNIESVELVLEE	1206
XX	AAW06417 standard; protein; 1169 AA.		RESULT 12			
XX	AAW06417		ID	AAW06417		
XX	AAW06417;		AC	AAW06417;		
XX	16-OCT-2003 (revised)		DT	16-OCT-2003		
XX	28-JAN-1997 (first entry)		DT	28-JAN-1997		
XX	Antiscarab pest toxin 50C(b).		DE	Antiscarab pest toxin 50C(b).		
XX	Toxin 50C(b); scarab pest; toxin; insect; scarabaeidae; pest control;		KW	Toxin 50C(b); scarab pest; toxin; insect; scarabaeidae; pest control;		
XX	larval stage insect; grain; tuberous crop; white grub; chafer grub;		KW	larval stage insect; grain; tuberous crop; white grub; chafer grub;		
XX	cyclocephala; popillia.		KW	cyclocephala; popillia.		
XX	Bacillus thuringiensis; strain kumamotoensis.		OS	Bacillus thuringiensis; strain kumamotoensis.		
XX	US5554534-A.		PN	US5554534-A.		

XX	Bacillus thuringiensis.	116	DB	415	VLLDIYVPGYVIFPGMPVEFPMVQLNNTK-----TLKYNFVSK	456
OS			QY	471	-----EIPPS--NEPLSRGSHRLSHITSYSFSKNASSPARYMGLPVPANTHRS	517
XX	WO200234774-A2.		DB	457	DIIASTRDSELELPETSQPNVYESYHRLCHITSIPATGNTT-----GLVPFVSWTHRS	511
PD	02-MAY-2002.		QY	518	ADVTNIVSDKLTQTPVKAHTLVSGTIVIKGPGTGEMILKETSS-----GPLAYTSVS	573
XX	24-OCT-2001; 2001WO-US045468.		DB	512	ADLNTIYSDKLTQTPAVRCMDNLPFPVVKPGHGTGGDLLQYNRSTGSGVTLFLARYGL	571
XX	24-OCT-2000; 2000US-0242838P.		QY	574	KSPLSORYRARIYASTTNLRFPVTISGTRIVSNVKNKWKGGDLTFTFIDLATIGTAF	633
PR	23-OCT-2001; 2001US-00032717.		DB	572	ALEKAKYRVRURYATDADIVLHND-----QIOMPKNPFGEDJTSKTFKVADAITL	626
XX	(DUPO) DU PONT DE NEMOURS & CO E I.		QY	634	---TFNSYSDSLTVGADSPAS--GGEVYVDKFLIPVNATFEABEDLDVAKAVNGLPSTK	689
PI	Abad AR, Duck NB, Feng X, Flannagan RD, Kahn TW, Sims LE;		DB	627	NLATDSSLAKENLGEDPNSTLSGIVYVDRIEFIPVDETVEABQDLAAKAVNALPTMT	686
XX	WPI: 2002-519178/55.		QY	690	KDALQTSVTDYVQNAANLVECLSDLYPNKXEMLDVAKEALVQARNLLQDTPGNRI	749
DR	N-PSDB; ABR67234.		DB	687	KDGLRPGVTDYEVNQNAANLVECLSDLYPNKXRLFLDVAKEALVQARNLLQDTPQEI	746
XX	New isolated pesticidal polypeptide useful for impacting insect pest e.g. Colorado potato beetle.		QY	750	NGENGWTSGTGLEVAEGDVLFXDRSLRLTSAREIDTETPTLYLQQIDESLLAPYTRYKL	809
PS	Claim 4; Page 96-98; 176pp; English.		DB	747	NGENGWTSGTGLEVAEGDVLFXDRSLRLTSAREIDTETPTLYLQQIDESLLAPYTRYKL	806
CC	The present invention relates to a new pesticidal polypeptide. The invention is useful for impacting an insect pest by applying the molecules of the invention to the environment of the insect pest by spraying, dusting, broadcasting, or seed coating, where the insect pest is selected from Colorado potato beetle, western corn rootworm or southern corn rootworm. The invention is also useful for increasing insect target range and for producing transgenic microorganisms and plants that express the pesticidal polypeptide. The invention is also useful for producing transformed plants and in transforming any organism to produce the pesticidal polypeptide of the invention. The present amino acid sequence represents a Bacillus thuringiensis wild-type Cry1218 endotoxin protein		QY	810	KGFIGSSQDLKILNHRANQIVKNVDPNLLPDLVFNKSGGIDRCSEQQYVDANALEN	869
CC			DB	807	RGFVSSQGLEIFTRHQTNRIVKNVDPNLLPDLVFNKSGGIDRCSEQQYVDANALEN	866
CC			QY	870	NGENGWSSDSHAFPHIDTGDIDNENGIWVFKIPTNGVATLGNLSEVEGPLSGE	929
CC			DB	867	R-----SGEHEFSIPIDTGEIDYENAGIWWGPKITDPEGVATLGNLSEVEGPLSGD	920
CC			QY	930	TLERAOQOQOQWQDKMKRGKSGAKYAAKQAI DRLFADYQDKLNSGVEMSLAAQN	989
CC			DB	921	ALERLQREQQWQKIQTRREEDRYMASKQAVDRLYADYQDQLNPDVEITDLTAQD	980
CC			QY	990	LVSQIPYVYNDALEIPGNYTSFTLTLNRLQOAWNLXDLRNAIPNGDFNGLSDWNATS	1049
CC			DB	981	LQISPIVYNEMFPEIPGMYTKFTLTLNRLQOAWNLXDLRNAIPNGDFNGLSDWNATS	1040
CC			QY	1050	DYNVQQLSDTSVLVFNKNSQVSOQFTVQPNRYLRTARKEGCGYVIRDGANQTE	1109
CC			DB	1041	GVEVQOQINETSVLVFNKNSQVSOQFTVQPNRYLRTARKEGCGYVIRDGANQTE	1100
CC			QY	1110	TLTENICDDDT--GVLSA-----	1125
CC			DB	1101	TUTFSASDITDGTNGYNTQVNTNGYNTNAYNTQASSTNGYNNANNMYNTQASNTGNTN	1160
CC			QY	1126	----DQTSYITTKTVEFTPESTEQVWIDMSETEGVFNIESVELVLEE	1167
CC			DB	1161	SVYNDQGTITTKTVPFIPYTDQMWIEMSETEGVFNIESVELVLEE	1206
XX	AAW06417 standard; protein; 1169 AA.		RESULT 12			
XX	AAW06417		ID	AAW06417		
XX	AAW06417;		AC	AAW06417;		
XX	16-OCT-2003 (revised)		DT	16-OCT-2003		
XX	28-JAN-1997 (first entry)		DT	28-JAN-1997		
XX	Antiscarab pest toxin 50C(b).		DE	Antiscarab pest toxin 50C(b).		
XX	Toxin 50C(b); scarab pest; toxin; insect; scarabaeidae; pest control;		KW	Toxin 50C(b); scarab pest; toxin; insect; scarabaeidae; pest control;		
XX	larval stage insect; grain; tuberous crop; white grub; chafer grub;		KW	larval stage insect; grain; tuberous crop; white grub; chafer grub;		
XX	cyclocephala; popillia.		KW	cyclocephala; popillia.		
XX	Bacillus thuringiensis; strain kumamotoensis.		OS	Bacillus thuringiensis; strain kumamotoensis.		
XX	US5554534-A.		PN	US5554534-A.		

QY	1	MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKYDILRMSEGENPELFGNP	60
DB	1	MSPNNQNEYIILDASSSTSVSDNSVRYPLANDQTTTLQNNYKYDILRMSEGENPELFGNP	60
QY	61	ETPIS--SSTVQIGIGVGLGALGVPPAGQIASFYSPIVQGLWPSSTVSVMKQVE	119
DB	61	EVLVSQDAKAADIVGKLLGLGVFPVGPVIVSLYQLIDILMPSGKSKSWEIPEQVE	120
QY	120	DLIDOKITDSVRKTALAGIQLGDLGVYQKSLKNWLENDRARSVVVTVYIALELDF	179
DB	121	ELINQKIARYARKALSELEGGLGNLYLTALKEEENPNSGRALDRVRNRFILDSLF	180
QY	180	VAKISFALSGORVPLLSVYAAANLHLRLDASIFGAWEFTPGIEISTFYDRQVTRTA	239
DB	181	QYMSFRTVNEVPLTVYMAANLHLRLDASIFGEWGSITTTINNYDRQVTRTA	240
QY	240	QYSDYCVKQWYNTGLDKLKTNAASWLKHYQFREMTLLVLDVALPNTYDTRPIETTA	299
DB	241	EYSDHCWKVYETGLAKLKTSAQWYDYNQFREMTLAVLDVALPNTYDTRPIETTA	300
QY	300	OLTREYVTDPIVFNRETSFGFCRWSLNSDISFSKESAVIRSPHLPDILSEIEPVTTRA	359
DB	301	OLTREYVTDPIVFNRETSFGFCRWSLNSDISFSKESAVIRSPHLPDILSEIEPVTTRA	356
QY	360	GLPLANTYLEYVGHVSHKYNKTNASSALERNYGTITS--NKIYYDLANKDIFQVRSLSGA	418
DB	357	S--ISSARIYRWAGHQISVHRVSRGSLNQWYGTQNLHSTSTFDTNVDYIKTSLKDA	414
QY	419	DLAN-----YYAQVYGVYASFTLLDK--NTSGSGVGFTYKPHHTMQVCTQNTYDID-	470

Query Match 53.7%; Score 3246.5; DB 5; Length 1206;
 Best Local Similarity 53.4%; Pred. No. 3 4e-205;
 Matches 665; Conservative 172; Mismatches 290; Indels 119; Gaps 18;

XX 10-SEP-1996.
 XX 30-SEP-1994; 94US-00315468.
 XX 16-DEC-1991; 91US-00808316.
 PR 30-JAN-1992; 92US-00828430.
 PR 01-FEB-1993; 93US-00014941.
 XX (MYCO) MYCOGEN CORP.
 XX Foncerrada L, Narva KE, Michaela TE;
 XX WPI; 1996-424659/42.
 DR N-PSDB; AAT43221.
 XX New nucleic acid encoding B.thuringiensis toxin active against scarab(s)
 PT - also related toxin and transformed microbes, effective against adult
 PT pests and their larvae.
 XX Claim 2; Col 29-36; 24pp; English.
 PS AAW06417-W06419 represent toxins that are active against scarab pests.
 XX The DNA encoding this sequence was isolated from the Bacillus
 CC thuringiensis strain kumamotoensis. Insects in the family Scarabaeidae
 CC constitute a serious pest control problem, especially when destructive
 CC larval stage insects infest high value turf found in golf courses,
 CC playing fields and lawns. The larvae of many species also attack grains,
 CC tuberous crops, and ornamentals. The larvae are known as white grubs, or
 CC chaffer grubs, and can be found in decaying organic matter, or in the soil
 CC where they consume plant roots. In Europe and the U.S. populations of
 CC these larvae and adults have developed resistance to chemical
 CC insecticides such as the organochlorines and DDT. These toxin sequences,
 CC and intact cells that are capable of expressing these proteins, can be
 CC used to control many pests of the family scarabaeidae, such as species of
 CC Cyclocephala, and Popillia. The toxins are active against larvae (present
 CC in soil) and against adults. (Updated on 16-OCT-2003 to standardise OS
 CC field)
 XX Sequence 1169 AA;
 Query Match 52.2%; Score 3152; DB 2; Length 1169;
 Best Local Similarity 54.0%; Pred. No. 5.7e-199;
 Matches 653; Conservative 179; Mismatches 294; Indels 84; Gaps 25;
 Qy 1 MSPNNQNEYILDASSSTVSNDNRYPLANDQTTTLQNNYKDYLRMSEGENPELFGNP 60
 Db 1 MSPNNQNEYIIDATPSTVSNDNRYPPANEPTNALQNMDDYLYKMSAGNVSEYPCGP 60
 Qy 61 ETPIS-SSTVQTGIGVQVGLGVPPAGQIASPYSFVGLWPSSTVSWEIMKQVE 119
 Db 61 EVFLSEQDAVKAADIVGKLTGLGVPPFVGVISLYTLQIDILWPSKQSQWEIPMEQVE 120
 Qy 120 DLIDQKITDVRKTAGLQGLDGLGVQKSLKNLENNDTRARSVVVTOYIALELDF 179
 Db 121 ELINQKIAEYARNKALSEGLGNNYQYLTALEENKPNKGRALRVRNRRLDLSLF 180
 Qy 180 VAKIPSPAISQGEVPLLSVYAAANLHLLLRDASIFGAEMGFTFGEISTFYDQVTRTA 239
 Db 181 TQYMPSPRVTVFVPLTVYTTAAANLHLLLRDASIFGEWGLSTSTNNYNNQMKLTA 240
 Qy 240 QYSDYCVKQWNTGLDKLGTAAWAKLKHQFRREMTLLVLDLWALFPNYDTRTYPIETTA 299
 Db 241 EYSDHCWKVYETGLAKLGSAAKQWIDYNQFRREMTLLVLDVVALFNSYDTRTYPLATTA 300
 Qy 300 QLTREYVTDPIVFNRETSQGCRRWSLNSDVSFSEVSAAVIRSPLHLDLSIEIRPYTTTRA 359
 Db 301 QLTREYVTDPIVFNRETSQGCRRWSLNSDVSFSEVSAAVIRSPLHLDLSIEIRPYTTTRA 356
 Qy 360 GLPLNTEYLYEYVWGHISIKYKNTNASSALERNYGT-----ITSNKIYDILANKDIFQVR 414
 Db 357 S--FTSDRYMYWAGHQISYKHIGTSSTFTQWYGTGNQLQSTN----FDTNVDYIKTL 410

Qy 415 SLGADLAN-----YAAQVGVPIASPTLLDKNTSGSGVGGFTYSKPHTMQVCTQNYNTI 469
 Db 411 SNGAVLLDIVPGYTYTTPFGMPETEPFMVQNNTRKT--LTY-KP-ASKQIIDRTROSE 466
 Qy 470 DEIPPE--NEPLSRGYSHRLSHITSYSPSKNASSPARYGNLPVPFAWTHRSADVNTVYSD 527
 Db 467 LELPETSGQPNYESYSHLGHIT-PIYSSSTST-----YVPVFSWTHRSADLTWTKSG 520
 Qy 528 KITQIPVVKAHTLVSGTGVVINGPGFTGNNILKRLTSSGPLAYTSVSVKGLSPLSORYAIRY 587
 Db 521 EITQIPGKSSITGRNTYIIKGRGYTGDLVALDR--IGSCFQMIIPESQRFIRIRY 578
 Qy 588 AS--TTNLRLEP-VTISGTRIYSINVKI-MKGD-DLTFTFDLATICTAFTFSNYSDSL 642
 Db 579 ASNETSYISLYGLNQSGT----LKFNTQYSKNENDLYND-----FYIETPRVI 625
 Qy 643 TVGADS-----FASGEVYV-DKFELIPVNATFEAREEDLDVAKAVNGLFTSKDAL 693
 Db 626 SVNASSNIQRLSIGIQTNWLPILDRISPIVDETYEARTDLEAAKAVNALPTTKDGL 685
 Qy 694 QTSVTDYQVNOAANLVECLSDLELYPNKRMWDVAKEAKRLVQANLLQDTGPNRINGEN 753
 Db 686 QPGVTDYEVNOAANLVECLSDLELYPNKRLLFDAVREAKRLSEARNLLQDDPQFQINGEN 745
 Qy 754 GWTGSGTGLEVAEGDVLKORSILRTSAREIDTETPTLYQQIDESLLKPYTRYKLKGP 813
 Db 746 GMTASTGIEVIEGDAVFKGRYLRPLPGAREIDTETPTLYQKVBEGVLKPYTRYLRGVP 805
 Qy 814 GSSQDLKELRHRANQIVKQNPONLLPDVLPVNSCGGIDRCSEQOQYVDANLALNNGEN 873
 Db 806 GSSQGLEIYTHQTNRIKQVPPDILLDPVPPVNDGRINRCSEQOQYVNSRLEVENR--- 862
 Qy 874 GMSDSDSHAFHIDTGBIDANENGTIWWPKIPITNGYATLGNLELVEEGPLSGTLER 933
 Db 863 ---SGEAHEPSIPIDTGEIDYNENAGIWWGPKITDPEGVATLGNLELVEEGPLSGDALER 919
 Qy 934 AQQCEQQQMDKARKGSEKAYAAKQAIIDRLFADYDOKLNSGVEMSDMLAQNLOS 993
 Db 920 LQKBEQWQKIQMTRRRETDTRYMAKQAVORLYADYQDQQLNPNVEITDLSAAQDLIOS 979
 Qy 994 IPYVNDALPEIPGNMYTSTFELTNRLQAAWNLVLRNAIPNGDFRNGLSDNATSDVNV 1053
 Db 980 IPYVYNEPPEIPGNMYTSTFELTNRLQAAWNLVLRNAIPNGDYRNLSNNNTTSGVNV 1039
 Qy 1054 QQLSDTSVLVTPNNSQVSQPTQPNRYVLRVYRTARKEGVDGTVIIRDGANQTHETLTP 1113
 Db 1040 QQINHTSVLVIPIPNWNEQVSQKFTVQPNQRYVLRVYRTARKEGVDGTVIIRDGANQTHETLTP 1099
 Qy 1114 NICDDDTGVLSDQTS-----YIT-----KTVEFTPTSTEQVWIDMSETGEVPI 1157
 Db 1100 SASDYDNGMYDTQASNTNGTNTNSVYMKPAISRKTVDISSVYNNQMWIEISTETGTFYI 1159
 Qy 1158 ESELVLLEEE 1167
 Db 1160 ESELVIDVE 1169
 RESULT 13
 AAW84581
 ID AAW84581 standard; protein; 1157 AA.
 XX AAW84581;
 XX 07-JUN-1999 (first entry)
 XX Amino acid sequence of Bacillus thuringiensis Cry9C mutant 15.
 XX Cry9C; crystal protein; toxic; mutant; insecticide; insect;
 XX transgenic plant; transgenic seed; lepidoptera.
 XX Bacillus thuringiensis.
 XX Synthetic.

PN W09900407-A2.
 XX
 PD 07-JAN-1999.
 XX
 PF 25-JUN-1998; 98WO-EP004033.
 XX
 PR 27-JUN-1997; 97US-00884389.
 XX
 PA (PLBZ) PLANT GENETIC SYSTEMS NV.
 XX
 PI De Roock S, Van Rie J;
 XX
 DR WPI; 1999-105666/09.
 XX
 XX New Bacillus thuringiensis Cry9C crystal proteins with improved toxicity
 PT - useful for producing transgenic plants in controlling pest insects,
 PT especially *Ostrinia nubilalis*, *Heliothis virescens* and *Diatraea*
 PT *grandiosella*.
 XX
 PS Example; Page; 40pp; English.
 XX
 CC This is the amino acid sequence of a Cry9C (crystal) mutant protein which
 CC is toxic to an insect species. It is modified in the method of the
 CC invention in an attempt to improve its toxicity. The improved proteins
 CC are useful for producing transgenic plants and seeds in controlling
 CC insects feeding on a plant. The transgenic plants are also useful for
 CC producing more transgenic plants with similar characteristics, or for
 CC introducing the improved cry9C gene (including fragments) into the same
 CC plant variety or related plant species. The Cry9C protein can be obtained
 CC from cultured cells of the transgenic plants and used in insecticide
 CC compositions against lepidopteran insects, especially *Ostrinia nubilalis*,
 CC *Heliothis virescens* and *Diatraea grandiosella*. The cry9C gene can be
 CC expressed in combination with another insect control protein (e.g. or
 CC another *B. thuringiensis* derived crystal protein, especially CryIb- or
 CC CryIIb-type protein), and used to transform *B. thuringiensis*, which
 CC produces other insecticidal toxins. Such strains are useful against a
 CC variety of insect pests, or insects where insect resistance development
 CC is prevented or delayed. (Note: this sequence is not given in the
 CC specification, but is generated using the information provided by the
 CC inventors.)
 XX
 SQ Sequence 1157 AA;

Query Match 51.8%; Score 3128; DB 2; Length 1157;
 Best Local Similarity 52.4%; Pred. No. 2.2e-197;
 Matches 620; Conservative 195; Mismatches 326; Indels 42; Gaps 14;

QY 1 MSPNNQNEYIILDASSSTVSNSRYPLANDQTTTLQNNYKYDYLMSGEPPELPGNP 60
 DB 1 MNRNNQNEYIILDAPHGCGPDDDDRYPLASDPNAALQNNYKYDYLQMTDEDYTDYINP 60
 QY 61 ETPISS-STVQGTGIGIVGVGLGALGVFPAGQIASFYSPIVGLWPSSTVSVMIMKQVE 119
 DB 61 SLISGRDAVQTALTVGVRILGALGVFPFGQIVSFYQFLTLTLPVNDTAIMAFPMQVE 120
 QY 120 DLIDQKITDSVRKTALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVYQVIALELDF 179
 DB 121 ELVYQKITEFARNQALRQLGLGDSFNVYQSLQNLWLADRDNTNLVSVRAQFIALLDLP 180
 QY 180 VAKIPSPAISGQVPLLSVTAQAANLHLLLRDASIFPAEWGFTPGBISTFYDRQVTRTA 239
 DB 181 VNAIPLFVAVNGQVPLLSVTAQAANLHLLLRDASIFPAEWGFTPGBISTFYDRQVTRTA 240
 QY 240 QYSDYCVKWTYNTGLDKLGTNNAKSLKVPFRREMTLLVLDLVALPNNYDTRTYPIETTA 299
 DB 241 KYTYCETWTNTGLDRGNTNTESWLYHOFRRMTLVLDVVALPNNYDTRTYPIETTA 300
 QY 300 QLTREVVYTDIVFNRRTSGGFCRRWSLNSDISFSSEVASIRSPHLPDIILSEIEFYTRA 359
 DB 301 QLTREVVYTDIVFNRRTSGGFCRRWSLNSDISFSSEVASIRSPHLPDIILSEIEFYTRA 359
 QY 360 GLPLNTEYLEYVWGHISIKYNTNASSALERNYGTITSNKIKY---YDLANKDIFQVRSI 416

DB 360 -FPV-SSNFMFYWSGHTLRSLNDSAVQEDSYGLITTTTRATINPGVDGTR-----IEST 413
 QY 417 GADLANYYAQVYGVVPVASFLL--LDKNTGSGVCGFTYSKPHTTMQVCTQYNTIDRIPP 474
 DB 414 AVDPASALIGIYGNRASFPVGGFLFNGCTTSPANGS-----CRDLYDNDLPP 461
 QY 475 ENBPLSRGYS-HRLSHITSYSPSKN-ASSPARYGNLFPVFAWTHRSADVNTVYSDKITOI 532
 DB 462 DE--STGSSTHRLSHVTPFPSPQNAQSIANAGSVPTVYVTRRDVLANITTPNRITQL 518
 QY 533 PVVKAHTLVSGTIVIKPGPFTGGNILKRTSGPLAYTSVSVKSPLSQRYRARIYASTTN 592
 DB 519 PLVKASAPVSGTIVLKGPGFTGGILRTTNGTFTGTLRTVTVNSPLTQQYRLRVFPASTGN 578
 QY 593 LRLPVTISGTRISYINVNKTMKGDDLTFTNTF---DLATIG---TATFSSNYSLSLTGA 646
 DB 579 FSTRVLGGVSGIDVRLGSTMARGOELTYESFPFRETTTTPGPPPTFTTQAQELITVA 638
 QY 647 DSPASGGEVYVDKPELIPVNAFPEAEEDLVAKAVNGLFTSKKDALQTSVTDYQVNOAA 706
 DB 639 EGVSTGGEYIDRIEIVPNPAREAREEDLEAAKAVASLFTKTRDGLQVNVTDYQVNOAA 698
 QY 707 NLYECLSDLEYPKEMKMLDVKAEKELVQARNLLQDTGNRING--RNGTGTSGTGIYA 764
 DB 699 NLVSCLSDEQYGDHKKMLLEAVRAAKRLSRERNLLQDPDFNTINSTENGWASGVITIS 758
 QY 765 EGDVLPKDRSLRLTSAREIDTETPTLYYQOIDBSLLKPYTRYKLKPGIGSSQDLLEIKLI 824
 DB 759 EGGPFFKGRALQASAR---ENYPTIYQKVDASVLKPYTRYRLDGPVKSQDLLEIKLI 814
 QY 825 RHANQIVKXVNPONLLPDVLFPVNSCGGIDRCSEQQYVDANLALNNGENGNGMSDSSHAPS 884
 DB 815 HHKHLVHLKXVNPONLVSDTYSDGSCGINRCDEQHQVDMQLDABHHPMDCCAAQTHEFS 874
 QY 885 FHIDTGEIDLNTGTVVWFKIPTTNGVATLGNLELVEEGSLGHTLERAQOQEQWODK 944
 DB 875 SYINTGLNASVDQGIWVLKVRTTDGYATLGNLELVEVGPLSGESLEREDRDNAKMAE 934
 QY 945 MAERKASERKAYTAAKQAIIDRLFADYQDQKLSNGVEMSDMLAAQNLVOSIPYVNDALPE 1004
 DB 935 LGRKRAEIDRVLAQAQAINHLFVDYQDQQLNPEIGLAINEASNLVESISGVYSDTLQ 994
 QY 1005 IPGNYTSTFELNRLQQAANLIDLRNAIPNGDFRNLSDWNATSDVNVQQLSDTSVLVI 1064
 DB 995 IPGINYEIYTELSDRLQAQASYLYTSRNAVQNGDFNSGLDSNNTTMDASVQOQGNHPLVL 1054
 QY 1065 PWNVSQVSOQFTQPNRYVLYRTARKEGVDGVYIIRDGANQTTETLTFTNCDDDTGVLS 1124
 DB 1055 SHWDAQVSOQLRVNPNCKYVLRVARKVGGDGYVTRDGAHQHTLTFNACDIDVNGTY 1114
 QY 1125 ADQTSYITKTVEFTPTTEQVWIDMSFTEGVFNIESVLELBE 1167
 DB 1115 VNDNSYITEEVFPETKHMVSESEGSFYIDSIETQOE 1157
 RESULT 14
 AAW84584
 ID AAW84584 standard; protein; 1157 AA.
 XX
 AC AAW84584;
 XX
 DT 07-JUN-1999 (first entry)
 XX
 DE Amino acid sequence of *Bacillus thuringiensis* Cry9C mutant 18.
 XX
 KW Cry9C; crystal protein; toxic; mutant; insecticide; insect;
 XX transgenic plant; transgenic seed; lepidoptera.
 OS *Bacillus thuringiensis*.
 XX
 XX *Bacillus thuringiensis*.
 XX
 PN W09900407-A2.

PD 07-JAN-1999.
 XX 25-JUN-1998; 98WO-BP004033.
 PR 27-JUN-1997; 97US-00884389.
 XX (PLBZ) PLANT GENETIC SYSTEMS NV.
 PA De Roock S, Van Rie J;
 PI WPI; 1999-105666/09.
 DR
 XX
 XX
 PT New Bacillus thuringiensis Cry9C crystal proteins with improved toxicity
 PT - useful for producing transgenic plants in controlling pest insects,
 PT especially *Ostrinia nubilalis*, *Heliothis virescens* and *Diatraea*
 PT *grandiosella*.
 XX
 XX Example; Page; 40pp; English.
 CC This is the amino acid sequence of a Cry9C (crystal) mutant protein which
 CC is toxic to an insect species. It is modified in the method of the
 CC invention in an attempt to improve its toxicity. The improved proteins
 CC are useful for producing transgenic plants and seeds in controlling
 CC insects feeding on a plant. The transgenic plants are also useful for
 CC producing more transgenic plants with similar characteristics, or for
 CC introducing the improved cry9C gene (including fragments) into the same
 CC plant variety or related plant species. The Cry9C protein can be obtained
 CC from cultured cells of the transgenic plants and used in insecticide
 CC compositions against lepidopteran insects, especially *Ostrinia nubilalis*,
 CC *Heliothis virescens* and *Diatraea grandiosella*. The Cry9C gene can be
 CC expressed in combination with another insect control protein (e.g.
 CC Another *B. thuringiensis* derived crystal protein, especially CryIAb- or
 CC CryIIb-type protein), and used to transform *B. thuringiensis*, which
 CC produces other insecticidal toxins. Such strains are useful against a
 CC variety of insect pests, or insects where insect resistance development
 CC is prevented or delayed. (Note: this sequence is not given in the
 CC specification, but is generated using the information provided by the
 CC inventors.)
 XX
 XX Sequence 1157 AA;

Query Match 51.8%; Score 3128; DB 2; Length 1157;
 Best Local Similarity 52.4%; Pred. NO. 2.2e-197;
 Matches 620; Conservative 195; Mismatches 326; Indels 42; Gaps 14;

QY 1 MSPNQNEIILDASSSTSVSNRYPLANDQTTTLQNNYKDYLRMSXENPELFGNP 60
 DB 1 MNRNNQNEIILDAFCGCPDSDVRYPLASDPNALQNNYKDYLRMSXENPELFGNP 60
 QY 61 ETPFIS-STVQIGIGVQVLCALGVPFAGQIASPYSFTVQGLWPSSTVSVWEMIMKQVE 119
 DB 61 SLSISGRDAVQATLVVGRILGALGVPPSGQIVSPYQPLNTLMPVNDTALWEAFMRQVE 120
 QY 120 DLIDQKITSVRKTLALQGLDGLDYVQKSLWLENDRTRSVVTVQVIALELDP 179
 DB 121 ELVNQOITEFAKQALQGLDGLDGFVYQSLQNLADNRNTRNLVRAQFIALDLDP 180
 QY 180 VAKIPFAISQGEVPLLSVYAAQNLHLILLDASIFGAEMGFTGCEISTFYDROWRTA 239
 DB 181 VNAIPLPAVQGOVPLLSVYAAQNLHLILLDASIFGAEMGFTGCEISTFYDROWRTA 240
 QY 240 QYSDYCVKMYNTGLDKLTNAASWLKHYQFREMFLVLDLVALFPNYDTYPIETTA 299
 DB 241 KYTNYCETWYNTGLDLRGNTESWLVYHQFREMFLVLDLVALFPNYDTYPIETTA 300
 QY 300 QLTREVYTDPIVNPRTSGCFRRWSLNSDISFSESVAIRSPHLPILSIEFYTTTA 359
 DB 301 QLTREVYTDPIVNPRTSGCFRRWSLNSDISFSESVAIRSPHLPILSIEFYTTTA 359
 QY 360 GLPLNNTLEYLVWGHISIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFQVRS 416
 DB 360 -FPV-SSNFMVWGHISIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFQVRS 416
 DB 360 -FPV-SSNFMVWGHISIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFQVRS 416

QY 417 GADLANYAQTIVGVYASFTL--LDKNTSGSGVGGTYTSKPHHTMQCTQNTYNTIDBIPP 474
 DB 414 AVDFRSALAGIYGVNRASFPVPGLEFNGTTSPPANGG-----CRDLTYDNDLPLP 461
 QY 475 ENEPLSRGYS-HRLSHITSYSFSKN-ASSPARVGNLPVPAWTHRSADVNTVYSDKITQI 532
 DB 462 DE---STGSSTHLSHVTFPSQTNQAGSIANGSGVPTVYVTRDVLANTIIIPNRIQL 518
 QY 533 PVVKAHTLVSGTVIKGPOFTGNIILKRTSSGLAYTSVSVKSPLSQRYPARIRYASTTN 592
 DB 519 PLVKASAPVSGITVLKGPFTGGILLRRTTNGTFTGLRTVNSPLAQOVLVRPASTGN 578
 QY 593 LRLFTVISTRIYSINNVNTMKGDDLTFTNP---DLATIG---TAFPTSNSDSLTVGA 646
 DB 579 PSIRVLRGVSGIDVRLGSTMNGOELIYESFTRETTTGPFPNPPFTTQAQEIITVNA 638
 QY 647 DSPASGEVYVDKPELIPVNAFTPEAREEDLDVAKKAVNGLFTSKDKALQTSVTDYQVNOAA 706
 DB 639 BGVSFGGEYIIDRIELVVPNPAREAREEDLEAKKAVASLFTTRDGLQNVNTDIYQVDOAA 698
 QY 707 NLVCECLSDLYPNEKRMMLMDAVKAEKRLVQARNLQDTGTFNRING--ENGWTSSTGIEVA 764
 DB 699 NLVCSLSDEQYCHDKKMLLEAVRAKRLSRERNLQDPDPFTNTINSTEENGKASNGVTIS 758
 QY 765 EGDVLFKORSRLTTSAREIDTFTYPTLYLQQIDESLLKPYTRYKLGFTGSSQDLKLI 824
 DB 759 EGGPFKGRALQASAR----ENYPTIYQVDASVLKPYTRYRLDGFVKSSQDLKLI 814
 QY 825 RHRANQIVKQVDPNLLPVLVPSNCGGIDRCSEQQVVDANLALENNGENGMSDSHAPS 884
 DB 815 HHKRVHLVKNVDPNLSVDTYSGSCSGINRCDEHQVQMDAHHFPMDCCEAAQTHEPS 874
 QY 885 PHIDTGEIDNENTGIVWVFKIPTTNGYATLGNLELVEEPLSGETLERAQQOQEQWQDK 944
 DB 875 SYINTGDLNASVDQGIWVVLKVRITDGYATLGNLELVEEPLSGESLERBQRDNKWNAE 934
 QY 945 MARKGASEKATYAAKQALDRLPADYQDQKLSGVEMSDMLAQNVLQSPYVYNDALPE 1004
 DB 935 LGRKEAREIDRVYLAQAQAINHLFVDYQDQQLAIPETGLAINEASNLVBSISGVYSDTLQ 994
 QY 1005 IFGMNTSFTELTNRLQQAANLYDLRNALPNCDFRNLGSDWNATSDVNWQQLSDTSVLVI 1064
 DB 995 IFGINYEIYELSDRLQQAASYLYTSRNAVQNGDFNSGLSSWNTTNDASVQDQGNMFLVL 1054
 QY 1065 PNWNSQVQOFTVQPNRYVRLVATARKGVGDGYVIRDGANGQTETLTENICDDDTGVL 1124
 DB 1055 SHWDAQVSQLRVNPNCKYVLRVATARKVGCGDGYVIRDGANGQTETLTENACDYNVTY 1114
 QY 1125 ADQTSYITKVEFTSTEQVWIDMSETEGVNIESVELVEEE 1167
 DB 1115 VNDNSYITEEVVFPETKHMWVEVSESGSFYIDSIFBIETOB 1157

RESULT 15
 AAW84571
 ID AAW84571 standard; protein; 1157 AA.
 XX
 AC AAW84571;
 XX
 DT 07-JUN-1999 (first entry)
 XX
 DE Amino acid sequence of *Bacillus thuringiensis* Cry9C mutant 5.
 XX
 XX Cry9C; crystal protein; toxic; mutant; insecticide; insect;
 KW transgenic plant; transgenic seed; lepidoptera.
 XX
 OS *Bacillus thuringiensis*.
 OS Synthetic.
 XX
 PN WO9900407-A2.
 PD 07-JAN-1999.
 XX

PF 25-JUN-1998; 98WO-EP004033.
 XX PR
 XX 27-JUN-1997; 97US-00884389.
 XX PA (PLBZ) PLANT GENETIC SYSTEMS NV.
 XX PI De Roeck S, Van Rie J;
 XX XX MPI; 1999-105666/09.
 XX DR
 XX PT New Bacillus thuringiensis Cry9C crystal proteins with improved toxicity
 PT - useful for producing transgenic plants in controlling pest insects,
 PT especially *Ostrinia nubilalis*, *Heliothis virescens* and *Diatraea*
 PT *grandiosella*.
 XX PS Example; Page; 40pp; English.
 XX CC This is the amino acid sequence of a Cry9C (crystal) mutant protein which
 CC is toxic to an insect species. It is modified in the method of the
 CC invention in an attempt to improve its toxicity. The improved proteins
 CC are useful for producing transgenic plants and seeds in controlling
 CC insects feeding on a plant. The transgenic plants are also useful for
 CC producing more transgenic plants with similar characteristics, or for
 CC introducing the improved cry9C gene (including fragments) into the same
 CC plant variety or related plant species. The Cry9C protein can be obtained
 CC from cultured cells of the transgenic plants and used in insecticide
 CC compositions against lepidopteran insects, especially *Ostrinia nubilalis*,
 CC *Heliothis virescens* and *Diatraea grandiosella*. The cry9C gene can be
 CC expressed in combination with another insect control protein (e.g.,
 CC Another *B.thuringiensis* derived crystal protein, especially CryIAb- or
 CC CryIIb-type protein), and used to transform *B.thuringiensis*, which
 CC produces other insecticidal toxins. Such strains are useful against a
 CC variety of insect pests, or insects where insect resistance development
 CC is prevented or delayed. (Note: this sequence is not given in the
 CC specification, but is generated using the information provided by the
 CC inventors.)
 XX CC
 XX SQ Sequence 1157 AA;

Query Match 51.78; Score 3126; DB 2; Length 1157;
 Best Local Similarity 52.38; Pred. No. 2.9e-157;
 Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

QY 1 MSPNNQMEYELDASSSTVSVDNSVRYPLANDQTTLLQNNYKDYLRMSGENSELPFCNP 60
 DB 1 MNRNNQMEYELIDAPHCQCFSDDDVRYPLASDPNAALQNNYKDYLRQWDBDYDTSYNP 60
 QY 61 ETPISS-STVQTGIGIVGVGLGALGVPPAGQIASFYFIVGQLWPSSTVSWEIMKQVE 119
 DB 61 SLGISGRDAVQVATLVVGRIGLGVPPFGQIVSFYQFLILTLVPPVNDTAIWEAPMRQVE 120
 QY 120 DLIDQKITDSVRKTALAGLGLGVLDVYQKSLKNWLENDRTRARSVVVTVQYIALSLDF 179
 DB 121 ELVNNQITFARNOALRQGLGDSFNVYQSRSLQNLARDNDTNLSVRAQFIALDLDF 180
 QY 180 VAKIPSPAISGQEVPLLSVYAQAANLHLLLRDASIFCAEWGFTPGBISTFYDRQVTRTA 239
 DB 181 VNAIPLFVNVNGQQVPLLSVYAQAANLHLLLRDASIFGEGNGFTQGHISTFYDRQELTA 240
 QY 240 QYSDYCVKWTNTGLDKLKGTAASWKLKHQFRREMTLLVLDLVALFPNYTRTYPIETTA 299
 DB 241 KYTNYCETWNTGLDRGTNTESWLRHYQFRREMTLVLDVVALFPYDYVRLYPTGNSP 300
 QY 300 QLTREVTYTDIVNRETPSGFCRRWSLNSDISSEVESAVIRSPHLEDLISBIEFYTTTRA 359
 DB 301 QLTREVTYTDIVNPPANAGLCRRWGTNPNTFSELENAFIRPHLPDLRLNSLTISNRR- 359
 QY 360 GLPLNNTLEYLYVGHGKIKYKNTVASSALERNYGTITSNKIKY---YDLANKOIFQVRSI 416
 DB 360 -FPV-SSNFMDYWSGHTLRSLYSLNDSAVQEDSYGLITTTTATINPGVDGTR----LEST 413
 QY 417 GADLANYAQQVGYVPYASFTL--LDKNVSGSGVGFTYKPHFTTMQVCTQNYNTIDRIPP 474

Db 414 AVDFRSALIGYGVNRASFVPGGLFNGTTSFANGG-----CRLDYDTNDELPP 461
 QY 475 ENRPLSRGYS-HRLSHITSYSFSKN-ASSPARYGNLVPFAWTHRSADVNTVYSKTIQOI 532
 Db 462 DE---STGSSTHRLSHVTFPSFQTNQAGSIANAGSVPTVWTRDVLNNTITPVRITQL 518
 QY 533 PVVKAMTLVSGTGVVIGKPGTGGNLLKRTSSGPLAYTSVSVKSPLSQRYRARIYASTTN 592
 Db 519 PLVKASAPVSGTGVVTLKPGTGGILRRRTTNGTFTLRTVTVNSPLTQYLRVRFPASTGN 578
 QY 593 LRLPVTISGTRIYSINVNKTNKGDLLTFNTF---DLATIG---TAPTFSNYSDSLTVCA 646
 Db 579 FSTRVLGGVSGIDVRLGSTMRCQELTYSPFTRRTTGTGFPNPPFTTQOABLTVNA 638
 QY 647 DSPAGSGEVYVDKPELIPVNAATFEAREDLVAKAVNGLPFTSKKDALQTSVTDVQVQAA 706
 Db 639 EGVSTGGEYIIDRIEIVPVNPARAREEDLEAAKAVASLPTKTRDGLQVNVTDVQDQA 698
 QY 707 NLVECLSDELYPNKRMKMDAVKEKELVQARNLLQDTGNRING--ENGWTSVGSIYA 764
 Db 699 NLVSCLSDEQYGHDKMLLEAVRAAKRLSRERNLLQDPDFTNTINSTEENGKASNGVTIS 758
 QY 765 EGDVLPDRSLRITSAREIDTETVPTLYQOIDEISLLKPYTRYKLGKIGSSQDLLEIKLI 824
 Db 759 EGGPFKGRALQLASAR----ENPTTYIKVYDASVLKPYTRYRLDGFVKSSQDLLEIDL 814
 QY 825 RHRANQIVKQVDPNMLPDLVLPVNSCGGIDRCSEQQYVDANLALNNGENGNNMSDSHAPS 884
 Db 815 HHKVVHLVKVDPNMLVSDTYSGSCGINRCDEQHQVDMQLDAEHPMDCCAAQTHBPS 874
 QY 885 PHIDTGBIDLNENTGIWVFKIPTNGYATLGNLELVEBGLSGETLERAQQOQEQWQDK 944
 Db 875 SYINTGDLNASVDQGIWVVLKVRTTDGYATLGNLELVEVGPLSGESLEREQRDNAKMNAB 934
 QY 945 MARKGASERKAYAAKQAIIDRLPADYQDQKLNKSGVEMSDMLAAQNLVQSIPIVYNDALPE 1004
 Db 935 LGRKRABIDRVYLAKQAIINHLFVDYQDQQLNPIGLAEINAEANLVEISICGVYSDTLAQ 994
 QY 1005 IPGNNTTSPTELNRLOQAAWNLVDLRNAIPNGDFRNGLSDNWATSDVNVQQLSDTSVLVI 1064
 Db 995 IPGINYEIYTELSDELQQAASYLYTSRNAVQMGDFNSGLDSWNTTMDASVQDQGNMFLVL 1054
 QY 1065 PNWNSQVSOQFTVQPNTRYVLRVTARKEGVGDGVYIIRDGANOTETLTNFICDDDTGVLS 1124
 Db 1055 SHWDAQVSOQLRVNPNCKYVLRVTARKVGGDGVYIIRDGANOTETLTNFACDDVDNGTY 1114
 QY 1125 ADQTSYITKTVETPSTPEQVWIDMSETEGVFNIESVELVLEEE 1167
 Db 1115 VNDNSYITEVWFPEKEMWVEVSESGSPYIDSIBFIETQB 1157

Search completed: June 21, 2004, 13:37:24

Job time : 73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:39:11 ; Search time 53 Seconds
(without alignments)
6216.227 Million cell updates/sec

Title: US-10-089-678-1
Perfect score: 6044
Sequence: 1 MSPNNQNEYILDASSSTVS.....MSETGVNIRSVELVLEE 1167

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	6044	100.0	1167	12	US-10-089-678-1
2	3263.5	54.0	1210	13	US-10-032-717-4
3	3263.5	54.0	1210	14	US-10-414-637-4
4	3263.5	54.0	1210	16	US-10-606-320-4
5	3246.5	53.7	1206	13	US-10-032-717-2
6	3246.5	53.7	1206	14	US-10-414-637-2
7	3246.5	53.7	1206	16	US-10-606-320-2
8	3100.5	51.3	1156	14	US-10-099-285-72
9	3100.5	51.3	1156	15	US-10-428-961-28
10	2884.5	47.7	1152	15	US-10-614-524-6
11	2825.5	46.7	1150	14	US-10-099-285-74
12	2710	44.8	1134	14	US-10-099-285-76
13	2519.5	41.7	1207	10	US-09-988-462-7
14	2457	40.7	1227	15	US-10-428-961-63
15	2454.5	40.6	1186	9	US-09-826-660-23

16	2449.5	40.5	1228	15	US-10-428-961-38	Sequence 38, Appl
17	2446.5	40.5	1228	15	US-10-614-524-2	Sequence 2, Appl
18	2230.5	36.9	1177	9	US-09-873-873-10	Sequence 10, Appl
19	2230.5	36.9	1177	9	US-09-873-873-12	Sequence 12, Appl
20	2230.5	36.9	1177	9	US-09-873-873-14	Sequence 14, Appl
21	2230.5	36.9	1177	10	US-09-997-914-10	Sequence 10, Appl
22	2230.5	36.9	1177	10	US-09-997-914-12	Sequence 12, Appl
23	2230.5	36.9	1177	10	US-09-997-914-14	Sequence 14, Appl
24	2230.5	36.9	1177	14	US-10-365-645-10	Sequence 10, Appl
25	2230.5	36.9	1177	14	US-10-365-645-12	Sequence 12, Appl
26	2230.5	36.9	1177	14	US-10-365-645-14	Sequence 14, Appl
27	2230.5	36.9	1177	16	US-10-672-163-10	Sequence 10, Appl
28	2230.5	36.9	1177	16	US-10-672-163-12	Sequence 12, Appl
29	2230.5	36.9	1177	16	US-10-672-163-14	Sequence 14, Appl
30	2229.5	36.9	1177	9	US-09-873-873-26	Sequence 26, Appl
31	2229.5	36.9	1177	10	US-09-997-914-26	Sequence 26, Appl
32	2229.5	36.9	1177	14	US-10-365-645-26	Sequence 26, Appl
33	2229.5	36.9	1177	16	US-10-672-163-26	Sequence 26, Appl
34	2227.5	36.9	1163	9	US-09-756-526A-2	Sequence 2, Appl
35	2227.5	36.9	1163	14	US-10-345-020-2	Sequence 2, Appl
36	2227.5	36.9	1163	14	US-10-342-821-2	Sequence 2, Appl
37	2225.5	36.8	1177	9	US-09-873-873-28	Sequence 28, Appl
38	2225.5	36.8	1177	10	US-09-997-914-28	Sequence 28, Appl
39	2225.5	36.8	1177	14	US-10-365-645-28	Sequence 28, Appl
40	2225.5	36.8	1177	16	US-10-672-163-28	Sequence 28, Appl
41	2215.5	36.7	1177	9	US-09-873-873-34	Sequence 34, Appl
42	2215.5	36.7	1177	14	US-10-365-645-34	Sequence 34, Appl
43	2191	36.3	1156	14	US-10-099-285-70	Sequence 70, Appl
44	2154.5	35.6	1174	9	US-09-826-660-6	Sequence 6, Appl
45	2154.5	35.6	1174	12	US-09-837-961-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-10-089-678-1
; Sequence 1, Application US/10089678
; Publication No. US20030017967A1
; GENERAL INFORMATION:
; APPLICANT: ASANO, Shinichiro
; TITLE OF INVENTION: PROTEIN HAVING INSECTICIDAL ACTIVITY, DNA ENCODING THE PROTEIN, AND METHOD
; FILE REFERENCE: 068821
; CURRENT APPLICATION NUMBER: US/10/089,678
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: JP 2000-236140
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: PCT/JP01/06660
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-089-678-1

Query Match	100.0%;	Score 6044;	DB 12;	Length 1167;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1167;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	MSPNNQNEYILDASSSTVS	SDNSVRYPLANDQTTT	LNMYKYLRMSGEENPELFGNP 60
Db	1	MSPNNQNEYILDASSSTVS	SDNSVRYPLANDQTTT	LNMYKYLRMSGEENPELFGNP 60
Qy	61	ETPISSTVGTGIVGQVILGALGV	PPAGQIASFVSPVTCGLWPSSTVS	VWEMIMKQVED 120
Db	61	ETPISSTVGTGIVGQVILGALGV	PPAGQIASFVSPVTCGLWPSSTVS	VWEMIMKQVED 120
Qy	121	LIDQKITDSVRKTKALAGLQGLG	LDVYQKSLKNLENRNDTRARS	VVVVTVYIALELDFV 180
Db	121	LIDQKITDSVRKTKALAGLQGLG	LDVYQKSLKNLENRNDTRARS	VVVVTVYIALELDFV 180

```
Qy 181 AKIPSPALSGQEVPLLSVYAAQANLHLLLDASIFGAENGFTPGESTFYDQVTRTAQ 240
Db 181 AKIPSPALSGQEVPLLSVYAAQANLHLLLDASIFGAENGFTPGESTFYDQVTRTAQ 240
Qy 241 YSDYCVKWTGLDKLKGTAASMLKXHQPREMTLLVLDLVALFPNYDTRTYPLETTAQ 300
Db 241 YSDYCVKWTGLDKLKGTAASMLKXHQPREMTLLVLDLVALFPNYDTRTYPLETTAQ 300
Qy 301 LTREYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLDILSBIETFYTRAG 360
Db 301 LTREYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLDILSBIETFYTRAG 360
Qy 361 LPLANTEYLEVWVGHISIKYKNTNASSALERNYGTITSNKIKYDILANKDIFQVRSLSGADL 420
Db 361 LPLANTEYLEVWVGHISIKYKNTNASSALERNYGTITSNKIKYDILANKDIFQVRSLSGADL 420
Qy 421 ANYAQAQVGVYPASFTLDDKNGSGVGGFTYKPHHTMQVCTQNYNTIDBIPENEPLS 480
Db 421 ANYAQAQVGVYPASFTLDDKNGSGVGGFTYKPHHTMQVCTQNYNTIDBIPENEPLS 480
Qy 481 RGYSHRLSHITSYSPSKNASSPARYGNLPVPAWTHRSADVTNTVYSDKITQIPVVKAHTL 540
Db 481 RGYSHRLSHITSYSPSKNASSPARYGNLPVPAWTHRSADVTNTVYSDKITQIPVVKAHTL 540
Qy 541 VSGTIVIKGPGFTGKNILKRTSSGPLAYTSVSVKSPLSQRYRARIYASTTNLRLFTVITS 600
Db 541 VSGTIVIKGPGFTGKNILKRTSSGPLAYTSVSVKSPLSQRYRARIYASTTNLRLFTVITS 600
Qy 601 GTRIYSINVKTMKNGDDLTPNTPLATIGTAPTSNYSDSITVGAQSFASGGEVYVDKP 660
Db 601 GTRIYSINVKTMKNGDDLTPNTPLATIGTAPTSNYSDSITVGAQSFASGGEVYVDKP 660
Qy 720 ELIPVNATFEAEEDLDVAKKAVNGIFTSKDALQTSVTDYQVQNAANLVECLSDLEYNE 780
Db 720 ELIPVNATFEAEEDLDVAKKAVNGIFTSKDALQTSVTDYQVQNAANLVECLSDLEYNE 780
Qy 781 KRMWLDAVKEAKRLVQANLLQDTGFNRINGENGWGTSGTIEVAEGDVLFKDRSLRLTSA 840
Db 781 KRMWLDAVKEAKRLVQANLLQDTGFNRINGENGWGTSGTIEVAEGDVLFKDRSLRLTSA 840
Qy 841 PDVLVNSCGGIDRCSEQYVDANLALENNNGENGMSSDSHAFSPHIDTGEIDLNENTGI 900
Db 841 PDVLVNSCGGIDRCSEQYVDANLALENNNGENGMSSDSHAFSPHIDTGEIDLNENTGI 900
Qy 901 WVVFKIPTNGYATLGNLLEVEGPLSGETTLERAQQQQOQOQODKMARXRGASEKAYYAAK 960
Db 901 WVVFKIPTNGYATLGNLLEVEGPLSGETTLERAQQQQOQOQODKMARXRGASEKAYYAAK 960
Qy 961 QAIIDLRFADYQDOKLNSGVENSDMLAQNLAQSIPIVYNDALPEIPGNANTSTFELTNRL 1020
Db 961 QAIIDLRFADYQDOKLNSGVENSDMLAQNLAQSIPIVYNDALPEIPGNANTSTFELTNRL 1020
Qy 1021 QQAANLYDLRNALPNGDFRNLGSDWNATSDVNVQQLSDITSVLVTPNNNSQVSOQFTVQPN 1080
Db 1021 QQAANLYDLRNALPNGDFRNLGSDWNATSDVNVQQLSDITSVLVTPNNNSQVSOQFTVQPN 1080
Qy 1081 YRYVLVTRARKEGVDGVVIRIDGANQTELTLPNI CDDDTGVLSDAQTSYITKTVEFTPS 1140
Db 1081 YRYVLVTRARKEGVDGVVIRIDGANQTELTLPNI CDDDTGVLSDAQTSYITKTVEFTPS 1140
Qy 1141 TEQWIDMSSETGVFNIESVELVLEEE 1167
Db 1141 TEQWIDMSSETGVFNIESVELVLEEE 1167
```

RESULT 2

US-10-032-717-4

; Sequence 4, Application US/10032717

```
; Publication No. US20020151709A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Peng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20020151709A1el Proteins With
; FILE OF INVENTION: Pesticidal Activity Against Coleopterans
; CURRENT APPLICATION NUMBER: US/10/032,717
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-10-032-717-4

Query Match 54.0%; Score 3263.5; DB 13; Length 1210;
Best Local Similarity 53.4%; Pred. No. 3e-258;
Matches 667; Conservative 178; Mismatches 284; Indels 119; Gaps 21;

Qy 1 MSPNQNEYEILDASSSTVSNSVRYPLANDQTTTLQNNMYKYOLRMSSEGENPELFGNP 60
Db 1 MSPNQNEYEILDASSSTVSNSVRYPLANDQTTTLQNNMYKYOLRMSSEGENPELFGNP 60
Qy 61 ETPFIS-STVQTGIGVQVLCALGVPPAGQIASYFIVGQMWSSITVVMEMKQVE 119
Db 61 ETVLSGQDAKAAADIVGKLLSGLGVPPVGVISLTQLIDILMFSGQKSQWEIPEQVE 120
Qy 120 DLIDQKITDSVRKALAGLQGLDGVYQSKNLWLNENDTRARSVVVTVIALELDF 179
Db 120 ELIQKLAIEYARNKALSEGLGNNYQLYLTALBEWKENPNSRALRDVNRFEILDSLF 180
Qy 180 VAKIPSPALSGQEVPLLSVYAAQANLHLLLDASIFGAENGFTPGESTFYDQVTRTA 239
Db 181 TOYMPSPFVTFVFPFLTVYTOAANLHLLLDASIFGAENGFTPGESTFYDQVTRTA 240
Qy 240 QYSDYCVKWTGLDKLKGTAASMLKXHQPREMTLLVLDLVALFPNYDTRTYPLETTA 299
Db 241 EYSDHCVKWYETGLAKLGTSAKQWVDYNQPREMTLLVLDLVALFPNYDTRTYPLETTA 300
Qy 300 QLTREYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLDILSBIETFYTR 359
Db 301 QLTREYVTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHLDILSBIETFYTR 356
Qy 356 GLPLANTEYLEVWVGHISIKYKNTNASSALERNYGTITS-NKIKYDILANKDIFQVRSLSGA 418
Db 357 S--ISSARYIRHMAGHQISYHRIPSDNIQKQYGTNQLHSTSTEDFTNYDIYKTLSDKA 414
Qy 419 DLAN-----YQAQVGVYPASFTLDDK--NTGSGSVGGFTYKPHHTMQVCTQNYNTID- 470
Db 415 VLLDITVFGYTYIIFGMEPEVFFMVNQLNLRK-----TLKKNPVSK 456
Qy 471 -----RIPPE--NEPLSRGYSRLSHITTSYSPSKNASSPARYGNLPVPAWTHRS 517
Db 457 DIIAGTRDSELELPETSDQPNYSYSRLCHITTSIPATGSTT-----GLVPFVSWTHRS 511
Qy 518 ADVNTVYSDKITQIPVVKAHTL---VSG-----TTVIKGPFTGGNILKRTSSGLPAYTSV 571
Db 512 ADLINAHVSDKITQIPVVKVSDLAPSTGGFPNNVTWVSGFGTGGGIKVRNG-VIISHM 570
Qy 572 SVK--SPLSQRYRARIYASTTNLRLFTVITSGTRIYSINVKTNKNGDDLTPNTFEDLATIG 630
Db 571 RVKISDINKYSMRIRYASANTTEFYINPSENVKS-HAQTKMNRGEALTNNKFNATL- 628
Qy 631 TAPTFSNYSDSLTVGADSFAS--GGEVYVDKFEIIPVNATFEAEEDLDVAKKAVNGLFT 687
```



```
Db 629 PPIKFTTTPPTLGAIFRABDFLGEAYIDRIEPIVDTEYAEODLEAKKAVNALFT 688
Qy 688 SKDALQTSVTDYQVNOAANLVECLSDLELYPNKRLMDAVKBAKRLVQARNLLQDTGFN 747
Db 689 NTKOGLRPGVTDYEVNOAANLVECLSDLLYPNEKRLLFDAVREAKRLSEARNLLQDPDFQ 748
Qy 748 RINGENGWSTGIEVABGVLPKDSRLRTSAREIDTETPTLYLYQOIDSLLKPYTRY 807
Db 749 EINGENGWSTAGTIEVIEGDALPKGYRLRPGAREIDTETPTLYLYQKVBEGVLKPYTRY 808
Qy 808 KLKGFICSSODLEIKLIRHRANOIYKVNPNOLLVDVLPVNSCGGIDRCSQQVVDANLAL 867
Db 809 RLKGFVGSQGLEIFTRHQTNRIYKVNPDLLPDPVSPVNSDGSINRCSQKYVNSLELV 868
Qy 868 ENNGENGWSSDSHAFSFIHDTGEIDLNENTGIWVVPKIPITTINGYATLGNLELVEGPLS 927
Db 869 ENR-----SGEARHFSIPIDTGEIDYNENAGIWWGFKITDPEGYATLGNLELVEGPLS 922
Qy 928 GETLERAQOQOQWQDMARKGASEKAYAAKQAIIDRLFADYQDQKLSGVEMSDMLAA 987
Db 923 GDALERLQREEQWKIQMTRREBETDRRYMASKQAVDRLYADYQDQQLNPDVEITDLTAA 982
Qy 988 QNLVQSIPIYVYNDALPEIPGMNTYSPTELNRLOQAAWNLVDLRNAIPNGDFRNGLSDNA 1047
Db 983 QDLIQSIPIYVYNNFPEIPGMNTYKFTELTDLRLOQAAWNLVDLRNAIPNGDFRNGLSDNA 1042
Qy 1048 TSDVNVQQLSDTSVLVTPNNSQVSOQFTVQPNRYVLRVTRARKEGVGDGYVIRDGANQ 1107
Db 1043 TPGVEVQOQIINHTSVLVPNWEQVSOQFTVQPNQRYVLRVTRARKEGVGNGVSIIRDGNQ 1102
Qy 1108 TETLTFNICDDDT-GVLASA----- 1125
Db 1103 TETLTFASDYDNGMYNTQVSTNGYNTNNAYNTQASTNGYNANNMYNTQASNTNGYN 1162
Qy 1126 -----DOTSYTKTWTZFTSTEQWIDMSETGCVNIESVELVLEE 1167
Db 1163 TNSVYNDQTYITKTVTFIPTYTQMWIEMSETGTFVIESVELIVOVE 1210

RESULT 3
US-10-414-637-4
; Sequence 4, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Peng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn B. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414,637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032,717
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-414-637-4

Query Match 54.0%; Score 3263.5; DB 14; Length 1210;
Best Local Similarity 53.4%; Pred. No. 3e-258;
Matches 667; Conservative 178; Mismatches 284; Indels 119; Gaps 21;

Qy 1 MSPNNQNEVEI IDATPSTSVSNDNRYPPANEPTNALQNMDYKDYLOMSAGNASEYPCSP 60
61 ETPIS-SSTVQTGIGIVGOVLGALGVPPFAGQIAFSYFIVGQLSPSSTVSVWEMIMQVE 119
61 EVLVSGQDAKAAAIIDIVGKLLSGLVFPVFGPIVSLYTQLIDILWPSQKQWEIFMEQVE 120
120 DLTDOKITDSVRKTALAGLQGLDGLDVYQKSLKNWLENENDTRARSVVVYQYTALELDF 179
121 BLINQKIAEYARNKALSELEGNGNVOYLTALEEMKENPNCGRALDRVRNRPHILDSLP 180
180 VAKIPSAISQGBVPLLSVYAQAANLHLRLDASIFGAEMGFTPGHISTFYDRQVTRTA 239
181 TQMPSEFVTFVFPFLVTVYQAANLHLRLDASIFGBEMGNTTITNNYDRQMKLTA 240
240 QYSDYCVKNTNTGLDKLKGNAASWLKHOFRBMTLLVLDLVALFPNYDTRTPIETTA 299
241 EYSDHCVKWYETGLAKLKGTSKQVNDYQFRBMTLLVLDVVALFPNYDTRTPIETTA 300
300 QLTREVVTDPVFNRETSQGFCCRRWSLNSDISFSEVASVIRSPLHLDLSEIEFYTTA 359
301 QLTREVVTDPVFNRETSQGFCCRRWSLNSDISFSEVASVIRSPLHLDLSEIEFYTTA 356
360 GLPLANNTEYLEYVWGHISIKYKNTNASSALERNYGTITS-NKIYYIDLANKDIQVRSIGA 418
357 S--ISSARYIRHWAGHQISYHRIEFDNI IKQYGTINQLHSTSTFDTNYDIYKLSKDA 414
419 DLAN-----YQAYQYGVYASFTLLDK--NTGSGSVGGFTYSKQHTTMQVCTQNTTID- 470
415 VLDDIVPPGYTYIFFGMEVEFPMVNLNTRK-----TLKYNPVSK 456
471 -----EYEPPE--NEPLSRGYSRLSHSTYSFSPKNASSPARGNLPVPAWTHRS 517
457 DIIAGTRDSLELPPEPSTSDQFNYESYSHRLCHITSIPATGSTT-----GLVPVFSWTHRS 511
518 ADVTNTVYSDKIQTQIPVVKAHTL---VSG---TWIKGPGTGGNMLKRTSSGPLAYTSV 571
512 ADLINAHSKIQTQIPVVKSDLAPSITGGPNNTVSGPGTGGIIVKIRNG-VIISHM 570
572 SVK-SPLSQRYARIRYASTNTNLRFLVTISGTRISYINVNKTKNGKDDLTFTNTPDLATIG 630
571 RVKISDINKYSNRIRYASANTTEFYINPSEENVKS-HAQTKNRGEALTYKNFYATL- 628
631 TAPTFSNYSLSLTVGADSPAS---CGEVYVDKPELIPVNAATFEAREDDLDVAKKAVNGLT 687
629 PPIKFTTTPPTLGAIFRABDFLGEAYIDRIEPIVDTEYAEODLEAKKAVNALFT 688
688 SKDALQTSVTDYQVNOAANLVECLSDLELYPNKRLMDAVKBAKRLVQARNLLQDTGFN 747
689 NTKOGLRPGVTDYEVNOAANLVECLSDLLYPNEKRLLFDAVREAKRLSEARNLLQDPDFQ 748
748 RINGENGWSTGIEVABGVLPKDSRLRTSAREIDTETPTLYLYQOIDSLLKPYTRY 807
749 EINGENGWSTAGTIEVIEGDALPKGYRLRPGAREIDTETPTLYLYQKVBEGVLKPYTRY 808
808 KLKGFICSSODLEIKLIRHRANOIYKVNPNOLLVDVLPVNSCGGIDRCSQQVVDANLAL 867
809 RLKGFVGSQGLEIFTRHQTNRIYKVNPDLLPDPVSPVNSDGSINRCSQKYVNSLELV 868
868 ENNGENGWSSDSHAFSFIHDTGEIDLNENTGIWVVPKIPITTINGYATLGNLELVEGPLS 927
869 ENR-----SGEARHFSIPIDTGEIDYNENAGIWWGFKITDPEGYATLGNLELVEGPLS 922
928 GETLERAQOQOQWQDMARKGASEKAYAAKQAIIDRLFADYQDQKLSGVEMSDMLAA 987
923 GDALERLQREEQWKIQMTRREBETDRRYMASKQAVDRLYADYQDQQLNPDVEITDLTAA 982
988 QNLVQSIPIYVYNDALPEIPGMNTYSPTELNRLOQAAWNLVDLRNAIPNGDFRNGLSDNA 1047
983 QDLIQSIPIYVYNNFPEIPGMNTYKFTELTDLRLOQAAWNLVDLRNAIPNGDFRNGLSDNA 1042
1048 TSDVNVQQLSDTSVLVTPNNSQVSOQFTVQPNRYVLRVTRARKEGVGDGYVIRDGANQ 1107
1043 TPGVEVQOQIINHTSVLVPNWEQVSOQFTVQPNQRYVLRVTRARKEGVGNGVSIIRDGNQ 1102
1108 TETLTFNICDDDT-GVLASA----- 1125
1103 TETLTFASDYDNGMYNTQVSTNGYNTNNAYNTQASTNGYNANNMYNTQASNTNGYN 1162
1126 -----DOTSYTKTWTZFTSTEQWIDMSETGCVNIESVELVLEE 1167
1163 TNSVYNDQTYITKTVTFIPTYTQMWIEMSETGTFVIESVELIVOVE 1210
```


Qy 1108 TETLTFNICDDDT-GVLSA----- 1125
Db 1103 TETLTFNSADYDTNGMTYQVNTNGYNTNAYNTQASSTNGYNANNKYNTOASNTNGYN 1162
Qy 1126 -----DOTSYITKTVFTPTSTEQWIDMSSETGFMIESVELVLEE 1167
Db 1163 TNSVYNDQGYITKTVTFIPYTDQWMIEMSETGTFVIESVELIVDVE 1210

RESULT 4
US-10-606-320-4
; Sequence 4, Application US/10606320
; Publication No. US20040091505A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Rafael Herrmann
; APPLICANT: Theodore W. Kahn
; APPLICANT: Albert L. Lu
; APPLICANT: Billy Fred McCutchen
; APPLICANT: James K. Presnail
; APPLICANT: James F.H. Wong
; APPLICANT: Cao-Guo Yu
; TITLE OF INVENTION: Genes Encoding Proteins With Pesticidal
; FILE OF INVENTION: Activity
; FILE REFERENCE: 35718/263948
; CURRENT APPLICATION NUMBER: US/10/606,320
; CURRENT FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: 60/391,786
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/460,787
; PRIOR FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 134
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1210
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-606-320-4

Query Match 54.0%; Score 3263.5; DB 16; Length 1210;

Best Local Similarity 53.4%; Pred. No. 3e-258;
Matches 667; Conservative 178; Mismatches 284; Indels 119; Gaps 21;

Qy 1 MSPNNQNEYELTDASSSTVSNDNRYPLANDQTTLLQNNYKDYLRMSGENPELFGNP 60
Db 1 MSPNNQNEYELTDATPSTVSNDNRYPPFANEPTNALQNMDDYKDYLRMSAGASEYFPGSP 60
Qy 61 ETPIS-SSTVQTGIGVQVLGALGVPPAGQIASFYSPIVQQLWPSSTVSVMIMKQVE 119
Db 61 EVLVSGQDAKAAIDVKGKLSGLGVPPGVPIVSLYTLQIDILWPSGKQWEIFMQVE 120
Qy 120 DLIPOKITDSVRKVALAGLOGDLGVQKSLKNWLRNDTRARSVVVTOYIALHLP 179
Db 121 ELINQKIAEYARNKALSELGLGNNYQLYLTALBEWKENPNSGRALDRVRNRPILDSLP 180
Qy 180 VAKIPSPAISQGVPLLSVYQAANLHLLIARDASIFGAEWGFTPGESTFYDQVTKTA 239
Db 181 TOTMPSFRVNTFVFPFLVYTOANLHLLIARDASIFGEWGSYTTINNYDRQMKLTA 240
Qy 240 QYSDYCVKQWNTGLDKLGTMAASWLKYHFRREMTLLVLDLVLPNNYDRTTYPIETTA 299
Db 241 EYSDHCWKVYETGLAKLGTSAKQWVDNQFRREMTLVLDVVALFPNNYDRTTYPMETKA 300
Qy 300 QLTREVTDPVFNRETSGGPCRWLSNDSISFEVESAVIRSPLFDILSEIFYTTRA 359
Db 301 QLTREVTDPVGLVAVNVSSIG---SW-YDKAPSPGVIESVIRPPHVFYITGLTVYTSR 356
Qy 360 GLPLNNTEYLYWVCHSISKYNTNASSALERNYGTITS-NKIYYDLANKDIPQVRSIGA 418
Db 357 S--ISSARYIRHWAGHQISYHRIIFSDNIIKQMYGTNQHLHSTFDFNTNYDIKTLSDKA 414

Qy 419 DLAN-----YAAQVYGVFYASFTLLDK--NTSGSGVGFTYSKPHHTMQVCTONTNID- 470
Db 415 VLLDIVPGYTYIFPGAEVEFPVWNLNTRK-----TLKNPVSK 456
Qy 471 -----EIPPE--NEPLSRGYSRLSHITSYSFSKNASSPARVGNLVPWATHRS 517
Db 457 DITAGTRDSELELPETSQDPNYSYSRLCHITSIPATGSTT-----GLVEVPFWSMTHRS 511
Qy 518 ADVTNTVYSDKI TOI PVVKAHTL---VSG---TTVIKPGFTGGMILKRTSSGPLAYTSV 571
Db 512 ADLINAHSDDKI TOI PVVKUSDLPASITGGPNVTVSGFGTGGGIKVIENG-VIISHM 570
Qy 572 SVK-SPLSORYRARIYASTTNLRLFTVITSGTRIYSINVNKMKMGDDLTPTNTPDLATIG 630
Db 571 RVKISDINKEYSMIRYASANTFEYINPSENVKS-HAQKTMNRGEALTYNKNFYATL- 628
Qy 631 TAPTFSNYSLSLTVGADSPAS---GGEVYVDKPLIPVNPATFEAEEDLDVAKAVNGHPT 687
Db 629 PPIKFTTTPFTITLGAIFEAEDPLGIEAYIDRIPIVDETYEAEQDLBAKCAVNAUPT 688
Qy 688 SKKDALQTSVTDYQVNOAANLVECLSDLYENKRLMDAVKRAKRLVQARNLLQDTGPN 747
Db 689 NTKDGLRPGVTDYEVNQANLVECLSDLYENKRLFLFDAVREAKELSEARNLLQDPPQ 748
Qy 748 RINGENGTGSGTIEVAEGDVLPKDRSLRLTSAREIDTETPTLYLQOISLLKPYTRY 807
Db 749 EINGENGTASTGIEVTEGDALPKGRVLRLLPGAREIDTETPTLYLQVKEGVLPKYTRY 808
Qy 808 XLKGPICSSQDLLEIKLRHANOIVKRVNPNLLPDVLPVNSCGGIDRCSEQYVDANLAL 867
Db 809 RLKGFVSSQGLEIFTIRHQTNRIVKRVDPDLLPDVSPVNSDGSINRCSEKYVNSRLLEV 868
Qy 868 ENNGENGNMSSDSHAFSHFIDTGEIDLNEMTGIWVVKPIPTTNGYATLGNLELVEEGPLS 927
Db 869 ENR-----SGEAHFSIPIDTGEIDYNENAGIWWGFKITDPEGVATLGNLELVEEGPLS 922
Qy 928 GETLERAQOQOQOQODKARRRGASEKAYAAKQAIIDLPAIDYQDKLNSGVMSDMLAA 987
Db 923 GDALERLQREQQWKIQWTRERRETDRRYMASQAVDRLYADYQDQQLNPPDVEITDLTAA 982
Qy 988 QNLVQSIPIYVYNDALPEI PCGNVYTSFTLTLNRLQOANVLYLRNAI PNGDFRGLSDWNA 1047
Db 983 QDLIQSIPIYVYNEMFPEI PCGNVYTKFTLTDRLQOANSLYDORNAI PNGDFRGLSNWNA 1042
Qy 1048 TSDVNVQQLSDTSVLVIVPNMNSQVSOQFTVQPNRYVLRVTRKEGVGDGVYIIRDGANQ 1107
Db 1043 TPGVEVQIINHTSVLAVIPNDEQVSOQFTVQPNQRYVLRVTRKEGVGVGVYIRDCGNQ 1102
Qy 1108 TETLTFNICDDDT-GVLSA----- 1125
Db 1103 TETLTFNSADYDTNGMTYQVNTNGYNTNAYNTQASSTNGYNANNKYNTOASNTNGYN 1162
Qy 1126 -----DOTSYITKTVFTPTSTEQWIDMSSETGFMIESVELVLEE 1167
Db 1163 TNSVYNDQGYITKTVTFIPYTDQWMIEMSETGTFVIESVELIVDVE 1210

RESULT 5

US-10-032-717-2

; Sequence 2, Application US/10032717

; Publication No. US20020151709A1

; GENERAL INFORMATION:

; APPLICANT: Andre R. Abad

; APPLICANT: Nicholas B. Duck

; APPLICANT: Xiang Peng

; APPLICANT: Ronald D. Flannagan

; APPLICANT: Theodore W. Kahn

; APPLICANT: Lynn B. Sims

; TITLE OF INVENTION: Genes Encoding No. US20020151709A1e1 Proteins With

; PESTICIDAL ACTIVITY AGAINST COLEOPTERANS

; FILE REFERENCE: 35718/237005

; CURRENT APPLICATION NUMBER: US/10/032,717

; CURRENT FILING DATE: 2001-10-23

```
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-032-717-2

Query Match      53.7%; Score 3246.5; DB 13; Length 1206;
Best Local Similarity 53.4%; Pred. No. 7.5e-257;
Matches 665; Conservative 172; Mismatches 290; Indels 119; Gaps 18;

QY 1 MSPNNQNEYIILDASSSTVSNSRYPLANDOTTTTLONNMYKDYLRMSGENPFLFCNP 60
DB 1 MSPNNQNEYIILDASSSTVSNSRYPLANDOTTTTLONNMYKDYLRMSGENPFLFCNP 60
QY 61 ETPIS-SSTVQTGIGIVGVGLGALGVPFAGQIASFYSPFVGLQWLPSTSVSWEMIMKQVE 119
DB 61 EVLVSGQDAKAALDIVGKLLSLGLVFPVGPVLSVLTQLDILWPSGEKSWEIFMEQVE 120
QY 120 DLIDQKITDSVRKTALAGLQGLDGVYQKSLKNWLENDRTRARSVVVTVQYIALELDP 179
DB 120 DLIDQKITDSVRKTALAGLQGLDGVYQKSLKNWLENDRTRARSVVVTVQYIALELDP 179
QY 121 ELINQKIABYARNKALSELEGLGNNYQLYLTALAEWEENPNSRALDRVNRFLDLSLF 180
DB 121 ELINQKIABYARNKALSELEGLGNNYQLYLTALAEWEENPNSRALDRVNRFLDLSLF 180
QY 180 VAKIPSPALSGQVPLLSVYAOAANLHLLLRDASIFGAEMGPTPGISTFYDROVTRTA 239
DB 180 VAKIPSPALSGQVPLLSVYAOAANLHLLLRDASIFGAEMGPTPGISTFYDROVTRTA 239
QY 181 TQMPSPFRVTNFPVPLTVYMAANLHLLLRDASIFGAEMGPTPGISTFYDROVTRTA 240
DB 181 TQMPSPFRVTNFPVPLTVYMAANLHLLLRDASIFGAEMGPTPGISTFYDROVTRTA 240
QY 240 QVSDYCVKMYNTGLOKLGKTNAAHSLKTHQFREMILLVLDLVALPNNYDTRTYPIETTA 299
DB 240 QVSDYCVKMYNTGLOKLGKTNAAHSLKTHQFREMILLVLDLVALPNNYDTRTYPIETTA 299
QY 241 EYSDHCWKVYETGLAKLGTSARQVYDQFREMILLVLDLVALPNNYDTRTYPIETTA 300
DB 241 EYSDHCWKVYETGLAKLGTSARQVYDQFREMILLVLDLVALPNNYDTRTYPIETTA 300
QY 300 QLTREYVYTOPIVNRRTSGCPERMNSLDSIFSEVESAVIRSPHFLDILSEIEFTTRA 359
DB 300 QLTREYVYTOPIVNRRTSGCPERMNSLDSIFSEVESAVIRSPHFLDILSEIEFTTRA 359
QY 301 QLTREYVYTOPIVNRRTSGCPERMNSLDSIFSEVESAVIRSPHFLDILSEIEFTTRA 356
DB 301 QLTREYVYTOPIVNRRTSGCPERMNSLDSIFSEVESAVIRSPHFLDILSEIEFTTRA 356
QY 360 GLPLNTEYLEYVWGHISIKYNTNASSALERNYGTITS-NKIYVYDLANKDIPQVSLGA 418
DB 360 GLPLNTEYLEYVWGHISIKYNTNASSALERNYGTITS-NKIYVYDLANKDIPQVSLGA 418
QY 357 S--ISSARIYRWAGHQSIVHRSRGSNLQMYGTQNLHSTSTPFTNVDYVYKLSKDA 414
DB 357 S--ISSARIYRWAGHQSIVHRSRGSNLQMYGTQNLHSTSTPFTNVDYVYKLSKDA 414
QY 419 DLAN-----YAAVYGVYASFTLLDK--NTGSGSGVGGFTYSKPHMTQMCTQNYATID- 470
DB 419 DLAN-----YAAVYGVYASFTLLDK--NTGSGSGVGGFTYSKPHMTQMCTQNYATID- 470
QY 415 VLIDVIVPGVYTFYFGKVEPEPMVQLNNTK-----TLKYNPVSK 456
DB 415 VLIDVIVPGVYTFYFGKVEPEPMVQLNNTK-----TLKYNPVSK 456
QY 471 -----RTPPE--NEPLSRGYSRHLSTYSFSCNASSPARYGNLVPATHRS 517
DB 471 -----RTPPE--NEPLSRGYSRHLSTYSFSCNASSPARYGNLVPATHRS 517
QY 457 DIIATRDSELEPLPETSQPNYSYSHRLCHITSIPATGNIT-----GLVPVFSWTHRS 511
DB 457 DIIATRDSELEPLPETSQPNYSYSHRLCHITSIPATGNIT-----GLVPVFSWTHRS 511
QY 518 ADVNTVYSDKIQTQIPVKAHLVSGTVTKGPGFTGNILKRTSS-----GPLAYTSVSU 573
DB 518 ADVNTVYSDKIQTQIPVKAHLVSGTVTKGPGFTGNILKRTSS-----GPLAYTSVSU 573
QY 512 ADLNTIYSDKIQTQIPVKAHLVSGTVTKGPGFTGNILKRTSS-----GPLAYTSVSU 571
DB 512 ADLNTIYSDKIQTQIPVKAHLVSGTVTKGPGFTGNILKRTSS-----GPLAYTSVSU 571
QY 574 KSPLSQRIRARIRASTTNRLPVTITSGTRIYSINVKTNKNGDDLTPTNFDTLATIGTAF 633
DB 574 KSPLSQRIRARIRASTTNRLPVTITSGTRIYSINVKTNKNGDDLTPTNFDTLATIGTAF 633
QY 572 ALEKAGKYAVLAYATDADIVLHVDA-----QIQPKTNPCEDLTSKTEKVAADITTL 626
DB 572 ALEKAGKYAVLAYATDADIVLHVDA-----QIQPKTNPCEDLTSKTEKVAADITTL 626
QY 634 ---TFPNYSDSLTVGADSPAS--GGEVYVDKFEIIPVNAIFEAEDLDVAKAINGLPTSK 689
DB 634 ---TFPNYSDSLTVGADSPAS--GGEVYVDKFEIIPVNAIFEAEDLDVAKAINGLPTSK 689
QY 627 NLATDSSALKRNLGDPNLSIGIYVVDRIEIPVDTEYAEQDLEAAKAVNALFTNT 686
DB 627 NLATDSSALKRNLGDPNLSIGIYVVDRIEIPVDTEYAEQDLEAAKAVNALFTNT 686
QY 690 KDALQTSVTDYQVNOQANLVECLSDLEYPNKCKMLMDAVKAEARLVQARNLLQDQFNRI 749
DB 690 KDALQTSVTDYQVNOQANLVECLSDLEYPNKCKMLMDAVKAEARLVQARNLLQDQFNRI 749
QY 687 KGLRGVTDYEVNOQANLVECLSDLEYPNKCKMLMDAVKAEARLVQARNLLQDQFNRI 746
DB 687 KGLRGVTDYEVNOQANLVECLSDLEYPNKCKMLMDAVKAEARLVQARNLLQDQFNRI 746
QY 750 NGENGWGTSGTIEVAEGDLVLFKDRSLRLTSARBIDTETPTLYLQYQIDESLLKPYTRYKL 809
DB 750 NGENGWGTSGTIEVAEGDLVLFKDRSLRLTSARBIDTETPTLYLQYQIDESLLKPYTRYKL 809
QY 747 NGENGWGTSGTIEVAEGDLVLFKDRSLRLTSARBIDTETPTLYLQYQIDESLLKPYTRYKL 806
DB 747 NGENGWGTSGTIEVAEGDLVLFKDRSLRLTSARBIDTETPTLYLQYQIDESLLKPYTRYKL 806
QY 810 KGFIGSSQDLIELKIRHRANQIVKQVNDLPLDVLVNSCGGIDRCSEQQYVDANLLEN 869
DB 810 KGFIGSSQDLIELKIRHRANQIVKQVNDLPLDVLVNSCGGIDRCSEQQYVDANLLEN 869
QY 807 RGPVSGSQGLEIPTIRHOTNRIYKQVNDLPLDVLVNSCGGIDRCSEQQYVDANLLEN 866
DB 807 RGPVSGSQGLEIPTIRHOTNRIYKQVNDLPLDVLVNSCGGIDRCSEQQYVDANLLEN 866
```

RESULT 6

```
US-10-414-637-2
; Sequence 2, Application US/10414637
; Publication No. US20030177528A1
; GENERAL INFORMATION:
; APPLICANT: Andre R. Abad
; APPLICANT: Nicholas B. Duck
; APPLICANT: Xiang Feng
; APPLICANT: Ronald D. Flannagan
; APPLICANT: Theodore W. Kahn
; APPLICANT: Lynn E. Sims
; TITLE OF INVENTION: Genes Encoding No. US20030177528A1el Proteins With
; TITLE OF INVENTION: Pesticidal Activity Against Coleopterans
; FILE REFERENCE: 35718/237005
; CURRENT APPLICATION NUMBER: US/10/414,637
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/10/032,717
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 60/242,838
; PRIOR FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1206
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-414-637-2
```

Query Match 53.7%; Score 3246.5; DB 14; Length 1206;
Best Local Similarity 53.4%; Pred. No. 7.5e-257;
Matches 665; Conservative 172; Mismatches 290; Indels 119; Gaps 18;

```
QY 1 MSPNNQNEYIILDASSSTVSNSRYPLANDOTTTTLONNMYKDYLRMSGENPFLFCNP 60
DB 1 MSPNNQNEYIILDASSSTVSNSRYPLANDOTTTTLONNMYKDYLRMSGENPFLFCNP 60
QY 61 ETPIS-SSTVQTGIGIVGVGLGALGVPFAGQIASFYSPFVGLQWLPSTSVSWEMIMKQVE 119
DB 61 EVLVSGQDAKAALDIVGKLLSLGLVFPVGPVLSVLTQLDILWPSGEKSWEIFMEQVE 120
QY 120 DLIDQKITDSVRKTALAGLQGLDGVYQKSLKNWLENDRTRARSVVVTVQYIALELDP 179
DB 120 DLIDQKITDSVRKTALAGLQGLDGVYQKSLKNWLENDRTRARSVVVTVQYIALELDP 179
QY 121 ELINQKIABYARNKALSELEGLGNNYQLYLTALAEWEENPNSRALDRVNRFLDLSLF 180
DB 121 ELINQKIABYARNKALSELEGLGNNYQLYLTALAEWEENPNSRALDRVNRFLDLSLF 180
QY 180 VAKIPSPALSGQVPLLSVYAOAANLHLLLRDASIFGAEMGPTPGISTFYDROVTRTA 239
DB 180 VAKIPSPALSGQVPLLSVYAOAANLHLLLRDASIFGAEMGPTPGISTFYDROVTRTA 239
QY 181 TQMPSPFRVTNFPVPLTVYMAANLHLLLRDASIFGAEMGPTPGISTFYDROVTRTA 240
DB 181 TQMPSPFRVTNFPVPLTVYMAANLHLLLRDASIFGAEMGPTPGISTFYDROVTRTA 240
```


Db 572 ALEKAGYRVRRLRYATDADIVLVHND-----QIQPKTWNPGEDLTKTFKADALITL 626
Qy 634 ---TFSSNYSDLTGADSPAS-GGEVYVVDKPELIPVNAITFEABEDLVAKKAVNGLFTSK 689
Db 627 NLATDSSLAKHNLGEPNSTLSIVVVDRIEPIPVDETTVEAQDLAKKAVNALFTNT 686
Qy 690 KDALQTSVTQVQNAANLECLSDLELYPNKRMMLDAVKEAKLVQARNLLQDTGNRI 749
Db 687 KDGLRPGVTDYEVNQAANLECLSDLLYPNEKRLLPDAVREAKRLSEARNLLQDPQEI 746
Qy 750 NGENGTGSGIIEVAEGDVLFKORSILTSAREIDTETPTLYLQOIDEILLKPYTRYKL 809
Db 747 NGENGTASTGIEVIEGDALFKGRLPLPGAREIDTETPTLYLQVKEGVLKPYTRYL 806
Qy 810 KGFPGSSQDLLEILIRHRAQIVKQVNDLLPDLVNSCGIDRCSEQQVVDANLLEN 869
Db 807 RGFVSSQGLEIFIRHQTWRIVKQVNDLLPDLVNSCGISNCRSEQKVNRSLEVEN 866
Qy 870 NGENGNSDSHAPSFHIDTGEIDLNERVTGIWVFKIPTNGYATLGNLELVBEGPLSGE 929
Db 867 R-----SGEHRFSIPIDTGEIDYENAGIWWGFKITDPEGYATLGNLELVBEGPLSGD 920
Qy 930 TLERAQOQEQOQDKWARKGASEKAYAAQADRLPADYQDOKLNSGVEMSDMLAQN 989
Db 921 ALERLQBEQOQKIQMTRRREBTRRRYMAKQAVDRLYADYQDQOLAPDVEITDLTAQD 980
Qy 990 LVQSIPIVYNDALPEIKQMYNYSFTELTLNRLQQAANLYDLRNLAPNGDFRNLSDMNATS 1049
Db 981 LIQSIPIVYNEPEIKQMYNYSFTELTLNRLQQAANLYDLRNLAPNGDFRNLSDMNATP 1040
Qy 1050 DVNVQQLSDTSVLVPIPNNSQVSOQFTVQPNRYRLVLTARKEGVDGQYVILRGANOTE 1109
Db 1041 GVEVQOINHSTSLVLPNWDQVSOFTVQPNRYRLVLTARKEGVDGQYVILRGANOTE 1100
Qy 1110 TLTFNICDDDT-GVLSA----- 1125
Db 1101 TLTFPSASDYDTNGMYNTQVNTNGYNTNAYNTQASSTNGYNNMNTQASNTNGYNTN 1160
Qy 1126 ----DQTSYTKTVEFPSTPEQWIDMSETEGVNIESVELMLEE 1167
Db 1161 SVYNDQGYTKTVPFPYTDQMMIENSETEGTYIESVELIVDVE 1206

RESULT 8

US-10-099-285-72

Sequence 72, Application US/10099285

Publication No. US20030105319A1

GENERAL INFORMATION:

APPLICANT: Schnepf, H. Ernest

Wicker, Carol

Narva, Kenneth E.

Waltz, Michelle

Stockhoff, Brian

Muller-Cohn, Judy

TITLE OF INVENTION: Toxins Active Against Pests

NUMBER OF SEQUENCES: 105

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/099,285

FILING DATE: 15-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/002,285
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 1156 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-099-285-72

Query Match 51.3%; Score 3100.5; DB 14; Length 1156;
Best Local Similarity 52.0%; Pred. No. 6.9e-245;
Matches 615; Conservative 196; Mismatches 330; Indels 41; Gaps 13;

Qy 1 MSPNNQNEYILDASSSTVSNSRYPLANDQTTTLQNNMYKYLRMSEGENPELFGNP 60
Db 1 MNRNNQNEYILDAHPCGCPSSDDVRPLASDPNAAALQNNMYKYLTQWDEDTDSYNP 60
Qy 61 ETPISS-STVQVIGIGVQVLGALGVPPAGQIASFYIVGQLMPSSTVSVMIMKQVE 119
Db 61 SLISGRDAVQTALTVVYILGALGVPSGQIVFYQFLLTLPVNDTALWEAFMRQVE 120
Qy 120 DLIDQKLTDSVRKTAGLQGLDGLVYQSKLNWLENRNDTRASVWVQYIALELDF 179
Db 121 ELVNQQTTFEARNQALRLAQGLGDSFNVYQSLQNLADRNDTRLSVVRQAQFIALDLD 180
Qy 180 VAKIPSPAISGQKVPYLLSVYQAANLHLLLRDASIFGAEMGFTFGEISTFVDQVTRTA 239
Db 181 VNALEPFAVNGQQVPLLSVYQAANLHLLLRDASIFGAEMGFTFGEISTFVDQVTRTA 240
Qy 240 QYSDYCVKQYNTGLDKLGTNAASWLKYHOPRREMTLLVLDLVALFPNYOTRTYDIETTA 299
Db 241 KYTNYCETWNTGLDLRGNTESWLYHQFRREMTLVLDVVALFPYVDVRLYPTGSPN 300
Qy 300 QLTREVTDDIVNRETSGGFCRRWSLNSDISFSEVESAVIRSHPHLPDILSIEPYYTRA 359
Db 301 QLTREVTDDIVNRETSGGFCRRWSLNSDISFSEVESAVIRSHPHLPDILSIEPYYTRA 359
Qy 360 GLPLNNTVEYLVVGHSHIKYKNTNASSALERNYGTITSNKIKY---VDLANKDIFQVRLS 416
Db 360 -FPV-SNFMDYSGHTLRSLYNDLSAQVEDSYGLITTRATINPGVDGTR-----IBST 413
Qy 417 GADLANYYAQVYGVYASFTL--LDKNTGSGVSGGFTYKSHPTTQVCTQYNTIDRIPP 474
Db 414 AVDFRSALIGYGVNRASFVPGGLFNGTTPANGG-----CRDLYDTNDELPP 461
Qy 475 ENBPLSRGYS-HRLSHITSYSPSKN-ASSPARYGNLPVPAWTHRSADVNTVYSDKITQI 532
Db 462 DB---STGSSTHRLSHVTFPFSPQTNQAGSTANAGSVPTVYVWTRDVLNNTITPNRI 518
Qy 533 PVVKAHTLVSGTIVKGFPGTGGNILKRTSGPLATYSVSKSPLSQRYRARIYASTTN 592
Db 519 PLVKASAPVSGTIVKGFPGTGGNILKRTSGPLATYSVSKSPLSQRYRARIYASTTN 578
Qy 593 LRLPVTISGTRIYSINVNKTMKGDLLTFNTFDLATTGT-----AFTPSYSDSLTVGAD 647
Db 579 PSIRILRGNTSIAYQRFSGTMRGQELTYESFVTSBPTTNGQSLPFTFTQAGNLTLAE 638
Qy 648 SPASGGGVYVDKPELIPVNAITFEABEDLVAKKAVNGLFTSKKDALQTSVTDYQVNOAN 707

```
Db 639 GVSTGSSYFIDRIEIIIPVNPAREEEDLEAAKCAVANLFTTRDGLQVNVTDYQVQDQAN 698
Qy 708 LVECLSDLYPNEKRLMDVAKEARLVQARNLQDTGFNRING--ENGWTSGTGIEVAB 765
Db 699 LVSCLSDEYQGHDKMLLEAVRAAKRLSRRENILQDDPDPNTINSTEENGWASNGVWISE 758
Qy 766 GDVLFKDRSLRLTSARBIDTETPTLYLYQOQIDESLLKPYTRYKLGPIGSSQDLLEIKLR 825
Db 759 GGPFFKGRALQALASAR---ENYPYIYQKVDAVSLKPYTRYLRDGFVKSSQDLLEIDLH 814
Qy 826 HRANQIVKXVNDNLDPVLPVNSCGGIDRCSEQQYVDANLALENNGENGMSDSSHAFSP 885
Db 815 HHKVLHVKXVNDNLVSDTYSDCSGINRCDEQHQVDMQDAEHHPMDCCAAQTHFSS 874
Qy 886 HIDTGEIDLNTENTGIVVVKLPITPTNGYATLGNLELVBEGLSGEETLERAAQOQOQWQDKM 945
Db 875 YINTGDINASVDQGIWVVKVTRTDGATLGNLELVBEGLSGEETLERAAQOQOQWQDKM 934
Qy 946 ARKGASEKAYAAQAIDRLFADYQDQKLGSGVMSDMLAAQNLVOSIPYVYNDALPEI 1005
Db 935 GRKEAIDRVYLAQAQINHLFVDYQDQALNPEIGLAEINBASNLVESISGVYSDTLQI 994
Qy 1006 PCMYTTSFTELNRLOQAWNLVDLNAIPNGDFRNLSDWNAATSDVNVQOLSSTSVLVIP 1065
Db 995 PGINYEIYTELSDRLQAQASLYLTSRNVAVQNGDFNSGLDSWNTTMDASVQDGNMHLVLIS 1054
Qy 1066 NNNSOVSOQFTVQPNRYRVLRTAKRKGVDGYVIRDCANOTETLTFNICDDDTGVLSA 1125
Db 1055 HWDQAQVSOQLRVNPKVLRVLRARVGGGJGYVIRDCAHQHETLTFNACDYVNGTYV 1114
Qy 1126 DQTSYITTKTEFTPTSTEQWIDMSETGVPNIESVELVLEE 1167
Db 1115 NDNSYITEEVVFPYETKHMVVESESGSFYIDSIEFTIETQE 1156

RESULT 9
US-10-428-961-28
; Sequence 28, Application US/10428961
; Publication No. US2003023711A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Kel
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rupar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; FILE REFERENCE: MECO201--1
; CURRENT APPLICATION NUMBER: US/10/428,961
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 28
; LENGTH: 1156
; TYPE: PRF
; ORGANISM: Bacillus thuringiensis
US-10-428-961-28

Query Match 51.3%; Score 3100.5; DB 15; Length 1156;
Best Local Similarity 52.0%; Pred. No. 6.9e-245;
Matches 615; Conservative 196; Mismatches 330; Indels 41; Gaps 13;

Qy 1 MSPNNQNEYIILDASSSTSVSNRYRPLANDOTTTLQNMNYKDYLRMSBEGNPELFGNP 60
Db 1 MNRNNQNEYIIDAFCGCPSSDDVRYPLASDPAALQNMNYKDYLMQTDYDTSYNP 60
Qy 61 ETFISS-STVQIGIGVQVIALGVPPAGQIATSPVSPVQOLWFPSSTVSVNEMIKQVE 119
```

```
Db 61 SLISGSDAVQVATLVVGRILGALGVPSGQIVSFPYQFLNLTLPVNDTALWEAFMRQVE 120
Qy 120 DLIDQKITDSVRKLTALAGLQGLDGLDLYVQSKLRNLENRNDTRARSVVVYQYIALBDF 179
Db 121 ELVNNQITTEPARNQALARLQGLGDFNVYQSLQNLADNRNLRNLVVVRAQFIALDLDF 180
Qy 180 VAKIPLSPALSGOEPHLLSVYAAQANLHLLLRDASIFGAEWGFTFGEISTSTFYDROVTRTA 239
Db 181 VNALPLFAVNGQQVPLLSVYAAQVNLHLLLRDASLFGEGWGTGEGISTSTFYDROVTRTA 240
Qy 240 QYSDYCVKRYNTGLDKLGTNAASWLKTHQPRREMTLLVLDLVALFPNRYDTRTYPIETTA 299
Db 241 KYTNYCETWNTGLDRAGTWTESWLRVHQPRREMTLLVLDLVALFPYDVELYPTGSPN 300
Qy 300 QLTREVTYDTPVNRBTSGGFCRRWSLNSDISPSEBSAVIRSRPHLPDILSIEBIETTRA 359
Db 301 QLTREVTYDTPVNRBPANVGLCRWGTNPYNTFSELENAPIRPHLPDLRLASLTSSNR- 359
Qy 360 GLPLNNTLEYLWVGHSTKYKNTNASSALERNYCTITSNKIKY---VDLANKDIFQVRS 416
Db 360 -FPV-SNFMWYSGHTLRRSYLNDASVQBSYGLITTRATINPVGDTNR-----IBST 413
Qy 417 GADLANYAAYVGVYASFTL--LDKNTGSGVGGFTYSKPHTTMQVCTQNTNTIDEIPP 474
Db 414 AVDFRSALIGIYGNRASFPVPGGLPNTTSPANGG-----CRDLYDTNDELPP 461
Qy 475 ENEPRLSRGYS--HRLSHITSYFSKN--ASSPARYGNLPVFAMTHRSADVNTVYSKITQI 532
Db 462 DE---STGSSTHRLSHVTFPFTQNGAGSIANAGSVPTVYVWTRRDVDLNNITPNRITQL 518
Qy 533 PVKATLTVSGTTVIKGPGFTGNLTKBTSSGLAYTSVSVKSPLSORVARIRVASTVN 592
Db 519 FLVKASAPVSGTIVLKGPGFTGGILRRTTNGTFTGLRVTVNSPLQTRVRVRFPASSGN 578
Qy 593 LRLFVTISGTRIYSINVNKNKGDLTFTNTFDLATIGT-----AFTFSNYSDSLTVGAD 647
Db 579 PSIRILRGNTSIAVQPCSTNWRGELTYESFTVTNQSDDLPTFTQAQENLTILAE 638
Qy 648 SPASGGEVYVDFKRLIPVNAITFRAEEDLDVAKAVNGLFTSKDQALQTSVTQYVQNAAN 707
Db 639 GVSTGSEYFIDRIEIIIPVNPAREEEDLEAAKCAVANLFTTRDGLQVNVTDYQVQDQAN 698
Qy 708 LVECLSDLYPNEKRLMDVAKEARLVQARNLQDTGFNRING--ENGWTSGTGIEVAB 765
Db 699 LVSCLSDEYQGHDKMLLEAVRAAKRLSRRENILQDDPDPNTINSTEENGWASNGVWISE 758
Qy 766 GDVLFKDRSLRLTSARBIDTETPTLYLYQOQIDESLLKPYTRYKLGPIGSSQDLLEIKLR 825
Db 759 GGPFFKGRALQALASAR---ENYPYIYQKVDAVSLKPYTRYLRDGFVKSSQDLLEIDLH 814
Qy 826 HRANQIVKXVNDNLDPVLPVNSCGGIDRCSEQQYVDANLALENNGENGMSDSSHAFSP 885
Db 815 HHKVLHVKXVNDNLVSDTYSDCSGINRCDEQHQVDMQDAEHHPMDCCAAQTHFSS 874
Qy 886 HIDTGEIDLNTENTGIVVVKLPITPTNGYATLGNLELVBEGLSGEETLERAAQOQOQWQDKM 945
Db 875 YINTGDINASVDQGIWVVKVTRTDGATLGNLELVBEGLSGEETLERAAQOQOQWQDKM 934
Qy 946 ARKGASEKAYAAQAIDRLFADYQDQKLGSGVMSDMLAAQNLVOSIPYVYNDALPEI 1005
Db 935 GRKEAIDRVYLAQAQINHLFVDYQDQALNPEIGLAEINBASNLVESISGVYSDTLQI 994
Qy 1006 PCMYTTSFTELNRLOQAWNLVDLNAIPNGDFRNLSDWNAATSDVNVQOLSSTSVLVIP 1065
Db 995 PGINYEIYTELSDRLQAQASLYLTSRNVAVQNGDFNSGLDSWNTTMDASVQDGNMHLVLIS 1054
Qy 1066 NNNSOVSOQFTVQPNRYRVLRTAKRKGVDGYVIRDCANOTETLTFNICDDDTGVLSA 1125
Db 1055 HWDQAQVSOQLRVNPKVLRVLRARVGGGJGYVIRDCAHQHETLTFNACDYVNGTYV 1114
Qy 1126 DQTSYITTKTEFTPTSTEQWIDMSETGVPNIESVELVLEE 1167
Db 1115 NDNSYITEEVVFPYETKHMVVESESGSFYIDSIEFTIETQE 1156
```


TELECOMMUNICATION INFORMATION:

TELEPHONE: (352) 375-8100

TELEFAX: (352) 372-5800

INFORMATION FOR SEQ ID NO: 76:

SEQUENCE CHARACTERISTICS:

LENGTH: 1134 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 76:

US-10-099-285-76

Query Match 44.8%; Score 2710; DB 14; Length 1134;
Best Local Similarity 48.1%; Pred. No. 7.9e-213;
Matches 559; Conservative 195; Mismatches 345; Indels 62; Gaps 19;

Qy 39 NMNYKDYLSRGENPL-----FGNPETPISSTVQTGIGVQVGLGVPFAGQIA 92
Db 4 NPNINECIPYCNLSPEVEVLGGERGNVT-----GLQTGIDIVAVVVGALGGVGGILT 58
Qy 93 SFYSPIVQLWPSSTVSWEMIMKQVEDLIDOKITDSVRKTALAGLQGLDLYQKSL 152
Db 59 GFLSTLPGFLWPSMDQAVFAFIQMEELIEQRISQOVVTRALDDLTGIONTYQYLIAL 118
Qy 153 KMWLENRDTRASVVVTVYIALEDFVAKIPSAISGQ-----EVPFLSVYAAANLHL 207
Db 119 KEMEERPNGVRA-NLVLRPEILLHALFVSSNPSPG-SGPGSQRFQALVYVYAAANLHL 176
Qy 208 LLLADASIPGAEGWTPGEBISTFTDROV-TRTAQSYDYCVKWTNTGLDKLGTWAAWLK 266
Db 177 LLLADAKEYGARWGLRESQIGNLYFNELOTRDTRDTYTHCVNAYNGLAGLGTSABSLWK 236
Qy 267 YHQPFRREMTLLDLVALFPNYDRTYPIETTAQLTRVYTDPIVFNKRTSGGP-----CR 322
Db 237 YHQPFRREMTLLDLVALFPNYDRTYPIETTAQLTRVYTDPIVFNKRTSGGP-----CR 322
Qy 323 RWSLNSDISFSEVASVIRSPHLPDILSEIFYTRAGLPLNNTFYLEYVWGHSI--KYK 380
Db 297 RMQTSAMTFNLNENALISSPHLPDITNNLMYTGSPFVHLTN-QLIEGMIGHSVTSLL 355
Qy 381 NTHASSALERNYGTITSNKIYDLANKDIPQRS-----LGADLANYYAQVGYVYAFYL 437
Db 356 ASGPTTVLLRNYGTTSS-IVNYFSFNDRVYQINTRSHTGLGFQV--APLFGITRAQF-- 410
Qy 438 LDKNMGSSVGGFTYKDPHTTMQVCTQNTYNTIDRIP--PENEPISRGVSHRLSHITSYSP 495
Db 411 -----YPGGTYSVTORNALTCBQYNSIDSLPDPNEPISRSYSHRLSHITSYSLH 461
Qy 496 -----SKNASSPARYGNLFPVANTHRSADVNTVYSDKTIQIPVVKAKHLVSGTTVWKG 549
Db 462 RVLITDGINIYS-----GNLPTVYVTHRDVDTLNTITADRIITQLPLVKSFPBAGTIVRG 517
Qy 550 PGFTGGMILKRTSGGPLAYTSVSKPLSQRYRARIYASTNLRPLFTTISGTRIYSINV 609
Db 518 PGFTGCDILRRTGVGTGTIRVTRTAPLQTRIRFRFASFTNPLFIRIGVGRDQVNYPDF 577
Qy 610 NKTANKGDDLTFNTFDLATIGTATFESNYDSLTVGADSPASGGEVYVDFKELLVNVATP 669
Db 578 GRVNRGDELRYESFATREBFTDIFNFRQQLISVFNANFSAGQEVTFDRILEIIPVAPAR 637
Qy 670 EABEDLVAKAVNGLFTSKDALQTSVTDYQVNAANLVESCLSDLYPNKRLMDAVK 729
Db 638 EAKEDLEAKAVASLFTRETDGLQVNVKQYQVQMAANLVESCLSDLYPNKRLMDAVK 697
Qy 730 EAKLVQARNLLQDTGNFRING--ENGWGTSGTGLEVAGDVLFDKRSRLTSAEIDTET 787
Db 698 AAKELSRERNLQDPDENTINSTENGWKAENGVTISGSGPFYKGRALQLASAR-----EN 753
Qy 788 YPTVLYQOIDSRLKPYTRYKLPFGIGSSQDLKLIHRANQIVKXVDPNLLPDVLPVN 847
Db 754 YPTVLYQKVDASELKPTRYRYSRSGDFVKSQDLKLIHRANQIVKXVDPNLLPDVLPVN 813

Qy 848 SCGGIDRCSEQQYVDANLALENNNGE-NGNMSDSDSHAPSFHIDTGEIDLNENGTIWWPKI 906
Db 814 SCGINRCQEQMYNAQLTEHHHPMDCCCAAQTHFESSYIDYDGLNSSVDQGIWAIKPV 873
Qy 907 PTTNGYATLGNLELVEGSLSGETLERAQQEQQWODKWKRGASAKYVYAAKQAIIDL 966
Db 874 RTTDGYATLGNLELVEGSLSGESLREQRONTKWSAELGRRAETDRVYQAKQSINHL 933
Qy 967 FADYQDQKLSGVMSDMLAAQNLVQSIPIYVYNDALPEIPGMNYTSTFTLTNRLQAAWNL 1026
Db 934 FVDYQDQALNPEIGCMADIMDAQNLVASISDVYSDAVLQIPGNIYIYITELSNRLQAQSVL 993
Qy 1027 YDLRNLAPNGDFRNLGSDMNATSDVNVQQLSDTSVLVLPNWNQSVSQOFTVQPNTRYVLR 1086
Db 994 YTSRNVONGDFNGLDSWNTAGASVQDQGNTHFLVLSHMDAQVSOQFRVQPNCKYVLR 1053
Qy 1087 VTAREKGVGVYIIRGANQCTETLTNIDDDTGVLSADOTSYITKTVBTPPSTEQWI 1146
Db 1054 VTAREKGVGVYIIRGANQCTETLTNIDDDTGVLSADOTSYITKTVBTPPSTEQWI 1146
Qy 1147 DMSETEGVFNIESVELVLEER 1167
Db 1114 EVNETGAFHIDSIEFVETEX 1134

RESULT 13

US-09-988-462-7

Sequence 7, Application US/09988462

Publication No. US20030046726A1

GENERAL INFORMATION:

APPLICANT: Koziel, Michael G.

Desai, Malini M.

Lewis, Kelly S.

Kramer, Vance C.

Warren, Gregory W.

Evola, Stephen V.

Crossland, Lytle D.

Wright, Martha S.

Merlin, Ellis J.

Lauris, Karen L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED

INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS:

ADDRESSEE: Syngenta Biotechnology, Inc.

STREET: 3054 Cornwallis Road

CITY: Research Triangle Park

STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/988,462

FILING DATE: 20-No. US20030046726A1-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/547,422

FILING DATE: 11-APR-2000

APPLICATION NUMBER: US 08/459,504

FILING DATE: 02-JUN-1995

APPLICATION NUMBER: US 07/951,715

FILING DATE: 25-SEP-1992

APPLICATION NUMBER: US 07/772,027

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: S-188051

TELECOMMUNICATION INFORMATION:


```

; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID
US-09-988-462-7

```

Query Match	41.7%	Score	2519.5	DB	10	Length	1207
Best Local Similarity	44.3%	Pred.	No. 3.9e-137				
Matches	546	Conservative	178	Mismatches	372	Indels	139
Gaps	20						
QY	37	LQNNYKDYLRMSGEGNELFGNPETFISSVTGTGIGVGOVLGALGVPPAGQIASFYS	96				
DB	4	LPDARIEDSLCIAG-----NWIDPFVASTVTGINIAGRIAGLVGVPPAGQIASFYS	57				
QY	97	FIVGQWPPSSVTSVWEMIKOVEDLIQKIDTSVRKLTALAGLGSLGDLVYQKSLKNWL	156				
DB	58	FLVGELWPRGR-DQWBIFFLEHVEQLINQOITENARNTALALQGLGSDSFRAYQOQSLDNL	116				
QY	157	ENRNDTPARSVVVTOYIALHLDVFAKIPSAISOQVPLLSVYAAQAAHTLLILLRDSIP	216				
DB	117	ENRNDARTSRVLYTOYIALHLDLFWNLFAIRNQEVPLFWYAAQAAHTLLILLRDSLF	176				
QY	217	GAEMGTFPGISIFYDRQVTRTAQSDYCVKQVNTGLDKLGTNAASWLKYHOFRRMTL	276				
DB	177	GSFGLTSQELQRYERQVTRDYSYCVEMVNTGLSLRGNTAASWVRYNQFRDLTL	236				
QY	277	LVLDLVALPNYDTRTPIFTTAQLTRVYTDPIVFNRTSGGFCRWSLNSDISFSFEVE	336				
DB	237	GVLDLVALFSSYDTRTPYINTSAULTREVYTDALGATGVNMAW-NWYNNAPSPSAIE	294				
QY	337	SAVIRSPLHFDLISEIFYTTRAGLPLNNTEYLEYVWGHISKYK-----NTNASSALER	390				
DB	295	AAAIRSHHLDLFLQLTIFS--ASSRWSNTRHMYWEGTHIQSRPIOGGLMSTHGATNT	352				
QY	391	NYGTITSNKIKYDILANKDIPQVRSGLADL--ANYAAQVYGVVPYASTLLDKNTGSGVG	448				
DB	353	SINPVTLR-----FASRDVVRVTSYAGVLLGWLYLPBHPVTRFVFNTPQNISDR-G	405				
QY	449	GPTYSKDHTMQVCTQNYNTIDELPPE--NEPLSRGYSHRLSHITSTSPSKNASSPARYG	506				
DB	406	TANYSQPYESPGQLKQDSEY--ELPPTTERPNYESHRLSHITGIILOQR-----V	455				
QY	507	NLPVFAWTHRSADVNTVYSDKIQIPIWVAHTLVSGTIVYIKGPGFGENILKRTSSGPL	566				
DB	456	NVPVYSWTHRSADRTWIGPNRIQIPIWVAASELPQGTTVVRGPGFGDILARTWTCGP	515				
QY	567	AYTSVSVKSPLSQRYARINRYASTNLRLPVITISGTRIYSINNVKNTWKGGDDLTFTNFDL	626				
DB	516	GPRIVTNGPLTKQYRIGFYASTVDPDFVSRGGTTVNNPRFLRTMNSDELKYGNFVR	575				
QY	627	ATTGATFSPNSDSLTVGADSPASGGEVYVDKFELIPVNATPEAREDLQVAKKAVNGLF	686				
DB	576	RAFTPTPTFOIQDIHRTSIGLSNGGEVIDKIEIIPVTATPEAYDLERAQAVNALP	635				
QY	687	TSKKD-ALQTSVDTVQVNOQAANLVECLSDBELYPNEKGMLMDVAKBAKRLVQARNLLQDTG	745				
DB	636	TNTNPRELKTVDVDYHDQVSNLVACLSDFECLDEKELKEKVKYAKRISDERNLLQDPN	695				
QY	746	FNRIIN-----GNGWGTSGTGIIVASGDVLFKORSRLRLTSAREIDT	785				
DB	696	FTSINKQPDFISTNEQSNFTSIHQSEHGHWGSENITIOEGNDVFKENYYTLPGT----FN	752				
QY	786	ETVPTYLYQOI DESLLAPYTRYKLKPGISSQDLLEIKLRHANOIWKVNP--DNLLPDV	843				
DB	753	ECTPTDYQKIGESLKATRYQLRGYIEUSQDLEIYLRYNAKHETLDVPGTESLWPLS	812				
QY	844	L--PVNSCGGIDRCSBQOYYDANLALNNKNGNG-NMSSDSHAFSFHIDTGEIDLNNTGI	900				

```

Db      813 VESPIGCEGPACRCA--PHPEWNPDLDCSDCEKCAHHSHHPSLSDIDVUCTDLHENLGV 870
Qy      901 WYVFKPTPTNGYATLGNLELIVRGPISGETLEBAQQOEOQWQDMAKRGASRKAYYAAK 960
Db      871 WYVFKLKTQBGHARLGNLFIEKPLLGEALSRVKAKEKWRDKRKLQLETKRVYTEAK 930
Qy      961 QATDRLPADYQDQKLSGVEMSDMLAAQNLVQSIPVYVNDALPRICGMVNTSFTELNRL 1020
Db      931 EAVDALFVDSQYDRLOQADTNMIGMHAADKLVRHIREAYLSLSELPVIFGVNARIPEELEGHI 990
Qy      1021 QQAWNLYDLRNLAI PNCD FRNGLSDMNATSDVNVQQLSDTSVLVIPNNNSQVSOQFTVPQN 1080
Db      991 ITAISLYDAERNVYKNGDFNNGLTCNVKGVHDVQQSHRSDI VIPKEBAEVSQAVRVCPG 1050
Qy      1081 YRVLRVTRAKGVGVDGYVIIRGANOTETLTP----- 1113
Db      1051 CGYILRVATYKEGYGECVVTIHEIENNTDELKPKNREBEVYPTDGTCDNYTAHQGTAG 1110
Qy      1114 -----NICDDDTGVLSAD--QTSY 1130
Db      1111 CADACNSRNAGYBDAYEVDPTASVNYKPTVBBEYTDVRRDNHCYDRGVNYPVPVPGY 1170
Qy      1131 ITRKIVFTPTSTEOVWIDMGSTEGVFNIESVELVLEBE 1167
Db      1171 VTKELSYFPETDVTWIEIGETGKPIVDSVELLMEE 1207

RESULT 14
US-10-428-961-63
; Sequence 63, Application US/10428961
; Publication No. US2003023711A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rugar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; FILE OF INVENTION: Polynucleotides, Compositions, and Methods of Use (Amend)
; FILE REFERENCE: MEC0201--1
; CURRENT APPLICATION NUMBER: US/10/428,961
; PRIOR FILING DATE: 2003-05-02
; CURRENT APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63
; LENGTH: 1227
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-428-961-63

```

Query Match	40.7%	Score 2457	DB 15	Length 1227
Best Local Similarity	42.8%	Pred. No. 5.4e-192		
Matches 541	Conservative 197	Mismatches 392	Indels 134	Gaps 20
Qy	1	MSPNNEVEYELDASSSTVSUNSVRYPLANDQTTLQNNMYKDYLRMGSEGENPELPGNP	60	
Db	1	LTSNRKNEEINALSIPAVNHSIAQWNLSTD-----ARIEDSLCIARG-----NNI	47	
Qy	61	ETPISSTVQTGIGVQWVLGALGVVPFAGQILASFYSFTVQGLPSPSTVSVWEMIMQVSD	120	
Db	48	DPFVSASTVQTGINIAGRIILSVLGVVPFAGQILASFYSFLVGELWPRGR-DPWEIFLBHVH	106	
Qy	121	LIDQKITDSVRKTPALAGLQGLGDLGVYQKSLKNWLENRNDTRARSVVVYQYLALELDFV	180	
Db	107	LIRQQTENTRTALARLQGLGNSFRAYQQSLDMLNRRDTRSVLVTQYLALELDFL	166	
Qy	181	AKIPSPAISQGEVPLLSVYAQAANTHLLLRDASIFGABGWFTPGISITFYDRQVTRTAQ	240	
Db	167	NAMPLPAIRNQEVPLLMVYAQAANTHLLLRDASLFGSGFGITQSIOIRYERQVKTRE	226	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2004, 13:36:05 ; Search time 23 Seconds
(without alignments)
2619.457 Million cell updates/sec

Title: US-10-089-678-1

Perfect score: 6044

Sequence: 1 MSPNNQNEYILDASSTVS.....MSTEGVFNIESVELVREE 1167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4007	66.3	1157	1	US-07-876-280-30
2	4007	66.3	1157	1	US-07-812-180A-2
3	4007	66.3	1157	1	US-08-315-468-2
4	4007	66.3	1157	3	US-07-941-650A-2
5	3662.5	60.6	1149	1	US-07-915-203-2
6	3662.5	60.6	1149	1	US-08-272-887-2
7	3662.5	60.6	1149	2	US-08-789-449-2
8	3152	52.2	1169	1	US-08-315-468-4
9	3123	51.7	1157	2	US-08-532-547-5
10	3123	51.7	1157	2	US-08-379-656B-5
11	3123	51.7	1157	3	US-08-455-838-5
12	3123	51.7	1157	3	US-09-019-809-5
13	3123	51.7	1157	4	US-09-471-177-5
14	3100.5	51.3	1156	4	US-09-002-285-72
15	3100.5	51.3	1156	4	US-09-589-477-72
16	3100.5	51.3	1156	4	US-09-661-322A-28
17	2825.5	46.7	1150	4	US-09-002-285-74
18	2825.5	46.7	1150	4	US-09-589-477-74
19	2715	44.9	1169	1	US-08-542-921-2
20	2715	44.9	1169	2	US-08-880-685-2
21	2715	44.9	1169	2	US-08-880-684-2
22	2710	44.8	1134	4	US-09-002-285-76
23	2710	44.8	1134	4	US-09-589-477-76
24	2549	42.2	1229	1	US-08-100-709-4
25	2549	42.2	1229	1	US-08-176-865-4
26	2549	42.2	1229	1	US-08-474-038-4
27	2549	42.2	1229	2	US-08-779-046-4

28 2549 42.2 1229 2 US-08-881-340-4 Sequence 4, Appli
29 2525 41.8 1227 3 US-09-053-549-2 Sequence 2, Appli
30 2519.5 41.7 1207 1 US-07-951-715A-7 Sequence 7, Appli
31 2519.5 41.7 1207 2 US-08-459-448A-7 Sequence 7, Appli
32 2519.5 41.7 1207 3 US-08-459-595A-7 Sequence 7, Appli
33 2519.5 41.7 1207 3 US-08-459-504B-7 Sequence 7, Appli
34 2519.5 41.7 1207 3 US-08-459-444-7 Sequence 7, Appli
35 2519.5 41.7 1207 3 US-09-053-549-8 Sequence 8, Appli
36 2519.5 41.7 1207 4 US-09-547-422-7 Sequence 7, Appli
37 2459 40.7 1227 1 US-08-448-170-8 Sequence 8, Appli
38 2459 40.7 1227 3 US-08-961-803-9 Sequence 9, Appli
39 2458.5 40.7 1138 1 US-07-973-320-2 Sequence 2, Appli
40 2457 40.7 1227 4 US-09-661-322A-63 Sequence 63, Appl
41 2454.5 40.6 1186 3 US-09-178-252-23 Sequence 23, Appl
42 2454.5 40.6 1186 4 US-09-826-660-23 Sequence 23, Appl
43 2452.5 40.6 1228 4 US-09-661-322A-38 Sequence 38, Appl
44 2441.5 40.4 1138 1 US-07-973-320-4 Sequence 4, Appli
45 2337.5 37.0 1177 3 US-08-754-490-14 Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-07-876-280-30
; Sequence 30, Application US/07876280
; Patent No. 5262158
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Bagley, Angela L.
; TITLE OF INVENTION: No. 5262158el Bacillus thuringiensis Isolates for
; TITLE OF INVENTION: Controlling Acarides
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/876,280
; FILING DATE: 19920430
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/S 104
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: PS50C
; IMMEDIATE SOURCE:
; CLONE: E. coli NM522(pMYC2320) NRRL B-18769
; US-07-876-280-30

Query Match 66.3%; Score 4007; DB 1; Length 1157;
Best Local Similarity 66.1%; Pred. No. 2.3e-285;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEYELDASSSTVSVDNSVRYPLANDQTTLLQNNYKDYLRMSKGENPFCNP 60
DB 1 MSPNNQNEYELDASSSTVSVDNSVRYPLANDQTTLLQNNYKDYLRMSKGENPFCNP 60

QY 61 ETFISSSTVOTGIGVGVGALGALGVPFAGQIASFYSTFVGLMPSSTVSVMEMKQVED 120
DB 61 ETFISSSTVOTGIGVGVGALGALGVPFAGQIASFYSTFVGLMPSSTVSVMEMKQVED 120

QY 121 LIDOKITDVKRTALAGLGGGLGVYQKSLKNWLENRNDTRARSVVVTVYIALELDV 180
DB 121 LIDOKITDVKRTALAGLGGGLGVYQKSLKNWLENRNDTRARSVVVTVYIALELDV 180

QY 181 AKIPSPASGOEVLSTVYAAANLHLLRLDASTFGAENGFTCEISTFYDROVTRTAQ 240
DB 181 AKIPSPASGOEVLSTVYAAANLHLLRLDASTFGAENGFTCEISTFYDROVTRTAQ 240

QY 241 YSDYCVKMYNTGLDKLKTGNAASMLKVHQPFRMTLLVLDLVALFPNYDRTYPIETTAQ 300
DB 241 YSDYCVKMYNTGLDKLKTGNAASMLKVHQPFRMTLLVLDLVALFPNYDRTYPIETTAQ 300

QY 301 LTRDYVTDPIAFNVTSTGFCNPMSTHSGILFYEVNNVIRPPLFDLSSVEINTSRGG 360
DB 301 LTRDYVTDPIAFNVTSTGFCNPMSTHSGILFYEVNNVIRPPLFDLSSVEINTSRGG 360

QY 361 LPIANNTVEYVWCHSICKYKNT-NASSALERNYGTITSNKIKYDVLANKDIPQVRSGLAD 419
DB 361 LPIANNTVEYVWCHSICKYKNT-NASSALERNYGTITSNKIKYDVLANKDIPQVRSGLAD 419

QY 420 LANYAQQVYGPYASFTLLDKNTGSGSVGGFTYKPTMTQVCTQNTYDIDEPENE-P 478
DB 420 LANYAQQVYGPYASFTLLDKNTGSGSVGGFTYKPTMTQVCTQNTYDIDEPENE-P 478

QY 478 LSRGVSHRLSHITSFSPKNSAPARYGNLVPVATHRASDVNTVYSDKTIQIPVVAH 538
DB 478 LSRGVSHRLSHITSFSPKNSAPARYGNLVPVATHRASDVNTVYSDKTIQIPVVAH 538

QY 538 VABSYSHRLSHITSFSPKNS--AYYGSFPVFWHTSADLNNITVYSDKTIQIPVAVRGD 535
DB 538 VABSYSHRLSHITSFSPKNS--AYYGSFPVFWHTSADLNNITVYSDKTIQIPVAVRGD 535

QY 598 TLVSGTIVIKGFTGNNILKRTSSGLAYTSVKSPLSORYARIYASVTLNLRPVT 598
DB 598 TLVSGTIVIKGFTGNNILKRTSSGLAYTSVKSPLSORYARIYASVTLNLRPVT 598

QY 598 MLYLGSVVGPGFTGGDILKRTNPSIIIGTAVTVNGSLQRVIRIYASTIDFE-FTL 594
DB 598 MLYLGSVVGPGFTGGDILKRTNPSIIIGTAVTVNGSLQRVIRIYASTIDFE-FTL 594

QY 599 ISGTRIYSINVKTNKGGDILTFATFDLATICTAFTFSYSDSLTVGADSFASGGEVYD 658
DB 599 ISGTRIYSINVKTNKGGDILTFATFDLATICTAFTFSYSDSLTVGADSFASGGEVYD 658

QY 654 YLGDITIEGNRFNKTMDNGASLTETFKFASFTDQFRETQDKILLSMGDFSSGQEVYD 654
DB 654 YLGDITIEGNRFNKTMDNGASLTETFKFASFTDQFRETQDKILLSMGDFSSGQEVYD 654

QY 659 KPELIPVNAVTEABEDLVAKAVNGLTFSKDALQTSVTDYVQVNOAANLVECLSDLYP 718
DB 659 KPELIPVNAVTEABEDLVAKAVNGLTFSKDALQTSVTDYVQVNOAANLVECLSDLYP 718

QY 714 RIEFIPVDVETYEAEQDLEAAKAVNALPNTKQGLRPGVTDYEVNOAANLVECLSDLYP 714
DB 714 RIEFIPVDVETYEAEQDLEAAKAVNALPNTKQGLRPGVTDYEVNOAANLVECLSDLYP 714

QY 719 NEKRLMDVAKAEKLVORNLQDTPGNRINENGWGTGTEIEVAGDVLFPKDRSLRT 778
DB 719 NEKRLMDVAKAEKLVORNLQDTPGNRINENGWGTGTEIEVAGDVLFPKDRSLRT 778

QY 774 NEKRLMDVAKAEKLVORNLQDTPGNRINENGWGTGTEIEVAGDVLFPKDRSLRT 774
DB 774 NEKRLMDVAKAEKLVORNLQDTPGNRINENGWGTGTEIEVAGDVLFPKDRSLRT 774

QY 779 SARREIDTETPTLYLQQIDESLLKPYTKLKGFTGSSQDLLEIKLIRHANOIVKNVDPN 838
DB 779 SARREIDTETPTLYLQQIDESLLKPYTKLKGFTGSSQDLLEIKLIRHANOIVKNVDPN 838

QY 834 GAREIDTETPTLYLQVKEGVKLPYTRYLRGFGVSSQGLEIYIRHQTNRIVKNVDPD 834
DB 834 GAREIDTETPTLYLQVKEGVKLPYTRYLRGFGVSSQGLEIYIRHQTNRIVKNVDPD 834

QY 839 LLPDVLVNSCGIDRCSEQQVVDANLALENNGENGNMSSDASHPSFHDITGEIDLNNT 898
DB 839 LLPDVLVNSCGIDRCSEQQVVDANLALENNGENGNMSSDASHPSFHDITGEIDLNNT 898

QY 888 LLPDVLVNSCGIDRCSEQQVVDANLALENNGENGNMSSDASHPSFHDITGEIDLNNT 888
DB 888 LLPDVLVNSCGIDRCSEQQVVDANLALENNGENGNMSSDASHPSFHDITGEIDLNNT 888

QY 898 GFWVFKIPPTNGYATLGNLLEVEGSLGSETLERAOQOQOQKMARKGASEKAYYA 958
DB 898 GFWVFKIPPTNGYATLGNLLEVEGSLGSETLERAOQOQOQKMARKGASEKAYYA 958

QY 948 GFWGFKITDPEGYATLGNLLEVEGSLGSETLERAOQOQOQKMARKGASEKAYYA 948
DB 948 GFWGFKITDPEGYATLGNLLEVEGSLGSETLERAOQOQOQKMARKGASEKAYYA 948

QY 959 AKQALDRLFADYQDQKNSGVMSDMLAAQNLVQISIPYVINDALPEIPGNNTSTFELTN 1018
DB 959 AKQALDRLFADYQDQKNSGVMSDMLAAQNLVQISIPYVINDALPEIPGNNTSTFELTN 1018

QY 949 SKQAVDLRYADYQDQQLNPDVEITDLTAQDLIQISIPYVYNEMFPEIPGNNTSTFELTD 1008
DB 949 SKQAVDLRYADYQDQQLNPDVEITDLTAQDLIQISIPYVYNEMFPEIPGNNTSTFELTD 1008

QY 1019 RLOQAWMLYDLRNLAPNGDFRGLSDWNAISDVNVQQLSDTSVLVPIPNWNSOVSOQFTVQ 1078
DB 1009 RLOQAWMLYDLRNLAPNGDFRGLSDWNAISDVNVQQLSDTSVLVPIPNWNSOVSOQFTVQ 1068

QY 1079 PNRYVLRVTRARKEGVGDDYVIRIDGANGQTLTIFNICODDGTGLSADQTSYITITVFT 1138
DB 1069 PNRYVLRVTRARKEGVGDDYVIRIDGANGQTLTIFNICODDGTGLSADQTSYITITVFT 1128

QY 1139 PSTQVWIDMSSETGVFNIESVBLVLEER 1167
DB 1129 PYTDQWMEISETGTFTVIESVBLVLDVE 1157

RESULT 2

US-07-812-180A-2
; Sequence 2, Application US/07812180A
; Patent No. 5366892
; GENERAL INFORMATION:
; APPLICANT: Focerrada, Luis R
; APPLICANT: Payne, Jewel M
; APPLICANT: Sick, August J
; TITLE OF INVENTION: No. 5366892el Coleopteran-Active Bacillus
; TITLE OF INVENTION: thuringiensis isolate and a No. 5366892el Gene Encoding a
; TITLE OF INVENTION: Coleopteran-Active Toxin
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Roman Saliwanchik
; STREET: 2421 N.W. 41st Street, Ste A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/812.180A
; APPLICATION NUMBER: 07/812.180A
; FILING DATE: 19920102
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, Roman
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: PS50C
; IMMEDIATE SOURCE:
; LIBRARY: Lambdagem (TM) - 11 LIBRARY OF LUIS
; LIBRARY: FONCERRADA
; CLONE: 50C
; US-07-812-180A-2

Query Match 66.3%; Score 4007; DB 1; Length 1157;
Best Local Similarity 66.1%; Pred. No. 2.3e-285;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEYELDASSSTVSVDNSVRYPLANDQTTLLQNNYKDYLRMSKGENPFCNP 60
DB 1 MSPNNQNEYELDASSSTVSVDNSVRYPLANDQTTLLQNNYKDYLRMSKGENPFCNP 60

QY 61 ETPFISSTVQTGTGIGVGVLCALGVPPAGQIASFYFIVGQLWPSSTVSVWEMIMKQVED 120
DB 61 ETPFISSTVQTGTGIGVGVLCALGVPPAGQIASFYFIVGQLWPSSTVSVWEMIMKQVED 120
QY 121 LIDOKITDTSVRKTAGLQGLGDLGVYQKSLKWLLENRNDTRARSVVVTOYIALSLDFV 180
DB 121 LVDQKLEKVKDRLALBELKGLGNALGVYQKSLKWLLENRNDTRARSVVVTOYIALSLDFV 180
QY 181 AKIPSPAGSQEVLPLSVYAQAANLHLLRLDASIFGAEMGPTPGHISTFYDQVTRTAQ 240
DB 181 SSIPSPAGSQEVLPLSVYAQAANLHLLRLDASIFGAEMGPTPGHISTFYDQVTRTAQ 240
QY 241 YSDYCVKWTNTGLDKLKTNAASNLKXHQPRREMTLLVLDLVALFPNYDTRTPPIETTAQ 300
DB 241 YSDYCVKWTNTGLDKLKTNAASNLKXHQPRREMTLLVLDLVALFPNYDTRTPPIETTAQ 300
QY 301 LTRVYDTPDIPVNETSGGFCRWSLNSDISFSEVESAVIRSPHLDILSEIEFPYTRAG 360
DB 301 LTRVYDTPDIPVNETSGGFCRWSLNSDISFSEVESAVIRSPHLDILSEIEFPYTRAG 360
QY 361 LPLNTEYLEYVWGHSHIKYKNT-NASSALERNYGTITSNKIKYDILANKDIPQVRLGAD 419
DB 361 LPLNTEYLEYVWGHSHIKYKNT-NASSALERNYGTITSNKIKYDILANKDIPQVRLGAD 419
QY 420 LANYIAQVGVPPASFTLLDKNTGSCSGVGGPTSKPHTTMQVCTQNYNIDIEIPPENE-P 478
DB 420 LANYIAQVGVPPASFTLLDKNTGSCSGVGGPTSKPHTTMQVCTQNYNIDIEIPPENE-P 478
QY 478 LKRGYSHRLSHITSYFSKNASPARVGNLPPAETHRSADVNTVYSDKIQTIPVVKAH 538
DB 478 LKRGYSHRLSHITSYFSKNASPARVGNLPPAETHRSADVNTVYSDKIQTIPVVKAH 538
QY 539 TLVSGTIVKPGPTGNTILKRTSGSLAVTSVSKPLSQRYRARRVASTTNLRLFT 598
DB 539 TLVSGTIVKPGPTGNTILKRTSGSLAVTSVSKPLSQRYRARRVASTTNLRLFT 598
QY 599 ISGTRISYINVTMKGDDLLFTNTFDLATIGTAFPSYSDSLTVGADSPASGGEVVD 658
DB 599 ISGTRISYINVTMKGDDLLFTNTFDLATIGTAFPSYSDSLTVGADSPASGGEVVD 658
QY 659 KEPLIPVATFARERDLDVAKAVALGLPTSKDALQTSVTDYOVNQAANLVECLSDRLYP 718
DB 659 KEPLIPVATFARERDLDVAKAVALGLPTSKDALQTSVTDYOVNQAANLVECLSDRLYP 718
QY 719 NEKRLMWDVAKRLVQARNLLQDTGPNRNGENGMTSGTIEVAGDGLFKDRSLRLT 778
DB 719 NEKRLMWDVAKRLVQARNLLQDTGPNRNGENGMTSGTIEVAGDGLFKDRSLRLT 778
QY 779 SAREIDTETPTLYQOIBESLLKPTRYKLKGFSSQDLKILRHRANQIVKQVDPN 838
DB 779 SAREIDTETPTLYQOIBESLLKPTRYKLKGFSSQDLKILRHRANQIVKQVDPN 838
QY 839 LLPDLVPSNCGGIDRCSCQVVDANLALENNENGMSSDSHAFSHIDTGTIDLENMT 898
DB 839 LLPDLVPSNCGGIDRCSCQVVDANLALENNENGMSSDSHAFSHIDTGTIDLENMT 898
QY 899 GIWVGFKIPTNGYATLGNLELVEEGLSETLRAQOQEQOQMDKQVRRKGRASERYA 958
DB 899 GIWVGFKIPTNGYATLGNLELVEEGLSETLRAQOQEQOQMDKQVRRKGRASERYA 958
QY 959 AKQAIORLADYQDKLNGVEMSDMLAQNVLQSIPIYVNDALPRIPGNVYTSFTELIN 1018
DB 959 AKQAIORLADYQDKLNGVEMSDMLAQNVLQSIPIYVNDALPRIPGNVYTSFTELIN 1018
QY 1019 RLQQAANLVDLRNAIPNGDFRNLGSDWNATSVDNVQQLSDTSVLVLPNWNQSVSQOFTVQ 1078
DB 1019 RLQQAANLVDLRNAIPNGDFRNLGSDWNATSVDNVQQLSDTSVLVLPNWNQSVSQOFTVQ 1078
QY 1079 PNRYVYLRVTAKEGVGDGVYIIRDGANQVTELTITFNICDDDTGVLSDAQTSYITKTVEFT 1138
DB 1079 PNRYVYLRVTAKEGVGDGVYIIRDGANQVTELTITFNICDDDTGVLSDAQTSYITKTVEFT 1138
QY 1139 PSTEQWIDMSETGEGVFNIESVELVLEEE 1167

DB 1129 PYTDQWIRIBSETGTFYIESVELVDVE 1157
RESULT 3
US-08-315-468-2
; Sequence 2, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Poncerra, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA: 07/808,316
; APPLICATION NUMBER: 16-DEC-1991
; FILING DATE: 16-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73.C2
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: PS50C
; IMMEDIATE SOURCE:
; LIBRARY: LambdaGen (TM) - 11 LIBRARY OF LUIS PONCERRADA
; CLONE: 50C(a)
; US-08-315-468-2
Query Match 66.3%; Score 4007; DB 1; Length 1157;
Best Local Similarity 66.1%; Pred. No. 2.3e-285;
Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;
QY 1 MSPNNQNEYILDASSSTVSNSVRYPLANDQTTLLQNNYKYDYLKMSGCEPPLFGNP 60
DB 1 MSPNNQNEYILDASSSTVSNSVRYPLANDQTTLLQNNYKYDYLKMSGCEPPLFGNP 60
QY 61 ETPFISSTVQTGTGIGVGVLCALGVPPAGQIASFYFIVGQLWPSSTVSVWEMIMKQVED 120

Db 1129 PYTDOMWIBISETGTFTYBSVELVDVE 1157

RESULT 4

US-07-941-650A-2

Sequence 2, Application US/07941650A

Patent No. 6294184

GENERAL INFORMATION:

APPLICANT: Oyeda, Kendrick A.

APPLICANT: Bradfisch, Gregory A.

TITLE OF INVENTION: Process for Controlling Lepidopteran Pests

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/07941.650A

APPLICATION NUMBER: US 07/941.650A

FILING DATE: 19920908

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/758,020

FILING DATE: 12-SEP-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/658,935

FILING DATE: 21-FEB-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/642,112

FILING DATE: 16-JAN-1991

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1157 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEetical: YES

ANTI-SENSE: NO

ORGANISM: Bacillus thuringiensis

STRAIN: kumamotoensis

INDIVIDUAL ISOLATE: PS50C

IMMEDIATE SOURCE:

CLONE: E. coli NM522 (pMYC1638), NRRL B-18751

US-07-941-650A-2

Query Match 66.3%; Score 4007; DB 3; Length 1157;

Best Local Similarity 66.1%; Pred. No. 2.3e-285;

Matches 773; Conservative 144; Mismatches 238; Indels 14; Gaps 7;

QY 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTLQNNYKDYLRMSBGENPBLFGNP 60

DB 1 MSPNNQNEYILDASSSTSVSDNSVRYPLANDQTTLQNNYKDYLRMSBGENPBLFGNP 60

QY 61 ETPFSSSTVGTGIGVGVILGALGVPVAGQIASFISFVGLWPSSTVSVWENKIMQVED 120

DB 61 ETPFSSSTVGTGIGVGVILGALGVPVAGQIASFISFVGLWPSSTVSVWENKIMQVED 120

QY 121 LIDQKITSVRKALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVYIALELDPV 180

Db 61 ETPFSSSTVGTGIGVGVILGALGVPVAGQIASFISFVGLWPSSTVSVWENKIMQVED 120

QY 121 LIDQKITSVRKALAGLQGLDGVYQKSLKNWLENRNDTRARSVVVTVYIALELDPV 180

Db 121 LVDQKIEKYVDKALAEKLGALDGVYQKSLKNWLENRNDTRARSVVVTVYIALELDPV 180

QY 181 AKIPSPALSGOEVPLLSVYAAQANLHLLLDASIPGAEWGTGEGISTFTFDQVTRTAQ 240

Db 181 SSIPSPAVSGHEVLLAVYAQVNLHLLLDASIPGAEWGTGEGISTFTFDQVTRTAQ 240

QY 241 YSDYCVKMYNGLDKLCTNAASLKYHOFERMTLLVLDVALPNNYDRTYPIETTAQ 300

Db 241 YSDYCVKMYNGLDKLCTNAASLKYHOFERMTLLVLDVALPNNYDRTYPIETTAQ 300

QY 301 LTRVYDTPDIPVNETSGGFCRRSLNSDISFSEVESAVIRPHLFDILSSVINTSRGG 360

Db 301 LTRVYDTPDIPVNETSGGFCRRSLNSDISFSEVESAVIRPHLFDILSSVINTSRGG 360

QY 361 LPLNTEYLEVWGHSLIKYKVT-NASSALERNYGTITSNKIKYDLANKOIPQVRSGLAD 419

Db 361 ITLANDAYINWSGHTLKRYRTADSTVYTYTANYGRITSEK-NSPALEDRIEINSTVAN 419

QY 420 LANYAQQVGPYASFTLLDKRTGSGVGGFTYSKPHHTMQVCTQNTYDIEIPENE-P 478

Db 420 LANYAQQVGPYASFTLLDKRTGSGVGGFTYSKPHHTMQVCTQNTYDIEIPENE-P 478

QY 478 LSRGYSRLSHITSYSPSKNASSPARYGNLVPFAMTHRSADVNTVYSDKLTQIPVVKAH 538

Db 478 VAESYSHLSHTSHSFKNGS--AYGSPFPVFWTHTSADLNTIYSDKLTQIPVVKAH 538

QY 539 TLVSGTIVKPGFTGGNKLKRTSGGLAYTSVSKPLSQRYRARIYASTTNLRLPVT 598

Db 539 TLVSGTIVKPGFTGGNKLKRTSGGLAYTSVSKPLSQRYRARIYASTTNLRLPVT 598

QY 598 MLYLGGSVVQGGFTGGNKLKRTSGGLAYTSVSKPLSQRYRARIYASTTNLRLPVT 598

Db 598 MLYLGGSVVQGGFTGGNKLKRTSGGLAYTSVSKPLSQRYRARIYASTTNLRLPVT 598

QY 599 ISGTRIYSINVKNTMKNKDDTLTPNFTDLATIGTFTPSNYSLSLTVGADSPASGEVVVD 658

Db 599 ISGTRIYSINVKNTMKNKDDTLTPNFTDLATIGTFTPSNYSLSLTVGADSPASGEVVVD 658

QY 658 YLGDITIEKNRNTKMDNGASITETFKFASFTIDPQFRETQDKILLSMGDFSSGQEVVD 654

Db 658 YLGDITIEKNRNTKMDNGASITETFKFASFTIDPQFRETQDKILLSMGDFSSGQEVVD 654

QY 718 KPELIPVNAPEAEEDLDVAKAVNGLFTSKKDALQTSVTDVQVNNQANLVECLSDLYP 718

Db 718 KPELIPVNAPEAEEDLDVAKAVNGLFTSKKDALQTSVTDVQVNNQANLVECLSDLYP 718

QY 714 RIEFIPVDETEAEQDLAAKAVNGLFTSKKDALQTSVTDVQVNNQANLVECLSDLYP 714

Db 714 RIEFIPVDETEAEQDLAAKAVNGLFTSKKDALQTSVTDVQVNNQANLVECLSDLYP 714

QY 719 NEKRLMDAVKAEKRLVQARNLQDTPNRRNGENGWGTGSGIEVAGDVLFKORSRLT 778

Db 719 NEKRLMDAVKAEKRLVQARNLQDTPNRRNGENGWGTGSGIEVAGDVLFKORSRLT 778

QY 778 NEKRLMDAVKAEKRLVQARNLQDTPNRRNGENGWGTGSGIEVAGDVLFKORSRLT 778

Db 778 NEKRLMDAVKAEKRLVQARNLQDTPNRRNGENGWGTGSGIEVAGDVLFKORSRLT 778

QY 838 SAREIDTETPTLYYQOQDESILKDYTRYKLGFTGSSQDLKILIRANQIVKNNVDDN 838

Db 838 SAREIDTETPTLYYQOQDESILKDYTRYKLGFTGSSQDLKILIRANQIVKNNVDDN 838

QY 834 GAREIDTETPTLYYQOQDESILKDYTRYKLGFTGSSQDLKILIRANQIVKNNVDDN 834

Db 834 GAREIDTETPTLYYQOQDESILKDYTRYKLGFTGSSQDLKILIRANQIVKNNVDDN 834

QY 898 LLPDLVPSNCGIDRCSEQVQVNDANLLENNGENGNSDPSHAPSPHIDTGEIDLNENT 898

Db 898 LLPDLVPSNCGIDRCSEQVQVNDANLLENNGENGNSDPSHAPSPHIDTGEIDLNENT 898

QY 888 LLPDLVPSNCGIDRCSEQVQVNDANLLENNGENGNSDPSHAPSPHIDTGEIDLNENT 888

Db 888 LLPDLVPSNCGIDRCSEQVQVNDANLLENNGENGNSDPSHAPSPHIDTGEIDLNENT 888

QY 958 GIWVFKIPTTNGYATLGNELVEGPISGETLERAOQOQOQOQOQOQOQOQOQOQOQOQ 958

Db 958 GIWVFKIPTTNGYATLGNELVEGPISGETLERAOQOQOQOQOQOQOQOQOQOQOQOQ 958

QY 948 GIWVFKIPTTNGYATLGNELVEGPISGETLERAOQOQOQOQOQOQOQOQOQOQOQOQ 948

Db 948 GIWVFKIPTTNGYATLGNELVEGPISGETLERAOQOQOQOQOQOQOQOQOQOQOQOQ 948

QY 1018 AKQALDRLPADYQOQ 1018

Db 1018 AKQALDRLPADYQOQ 1018

QY 1008 SKQAVDRLYADYQOQ 1008

Db 1008 SKQAVDRLYADYQOQ 1008

QY 1078 RLQQAANLYDLRNLAPNGDPRNGLSDMNATSDVNVQQLSDTSVLPVNNNSQVSOQFTVQ 1078

Db 1078 RLQQAANLYDLRNLAPNGDPRNGLSDMNATSDVNVQQLSDTSVLPVNNNSQVSOQFTVQ 1078

QY 1068 RLQQAANLYDLRNLAPNGDPRNGLSDMNATSDVNVQQLSDTSVLPVNNNSQVSOQFTVQ 1068

Db 1068 RLQQAANLYDLRNLAPNGDPRNGLSDMNATSDVNVQQLSDTSVLPVNNNSQVSOQFTVQ 1068

QY 1138 PNRYVLRVATARKEGVGDYVIRGANGOTETLTPNI CDDTGVLSADQTSYITIKVEPT 1138

Db 1138 PNRYVLRVATARKEGVGDYVIRGANGOTETLTPNI CDDTGVLSADQTSYITIKVEPT 1138

QY 1128 PNQVLRVATARKEGVGDYVIRGANGOTETLTPNI CDDTGVLSADQTSYITIKVEPT 1128

Db 1128 PNQVLRVATARKEGVGDYVIRGANGOTETLTPNI CDDTGVLSADQTSYITIKVEPT 1128

QY 1139 PSTEQVMDMSETEGVNIESVELVLEE 1167

Db 1139 PSTEQVMDMSETEGVNIESVELVLEE 1167

```
Db 121 LVQKLEKVKDQKALAEKGLGNALDVYQOSLEWLENRNDATRSVSNQFALDLPV 180
Qy 181 AKIPSPAIQORVPLLSVYAQAANLHLLLRDASIFGAEWGPTTGEISFYDRQVTRTAQ 240
Db 181 SSIPSPAVSGHEVLLAVYAQAANLHLLLRDASIFGAEWGPTTGEISFYDRQVTRTAQ 240
Qy 241 YSDYCVKVTYGLDKLGTNAASLKYHOPRREMTLLVLDLVALFPNYDTRTYPIETTAQ 300
Db 241 YSDYCVKVTYGLDKLGTNAASLKYHOPRREMTLLVLDLVALFPNYDTRTYPIETTAQ 300
Qy 301 LTRVETDPTVFNRETSGGFCRWSLNSDISFSEVESAVIRSPHLPDILSEIEFYFTRAG 360
Db 301 LTRVETDPTVFNRETSGGFCRWSLNSDISFSEVESAVIRSPHLPDILSEIEFYFTRAG 360
Qy 361 LPLNTEYLEYVWGHISIKYQNT-NASSALERNTGTTISNKIKYDLANKDIFPVRSLGAD 419
Db 361 ITLNDAYINYSGHTLKRYRTADSTVYTYANYGRITSEK-NSPALEDRLDIFRINSTVAN 419
Qy 420 LANYAQUVGPVAPSTLLDKNTGSGVGFTYSKPHTTMQVCTQNYNTIDIEPPENE-P 478
Db 420 LANYAQUVGPVAPSTLLDKNTGSGVGFTYSKPHTTMQVCTQNYNTIDIEPPENE-P 478
Qy 479 LRGYSRHLSHITSYSPKSSNASSPARYGNLPVANTHRSAADVNTVYSDKIQTIPVVKAH 538
Db 478 VAESYSHRSHITSYSPKSSNASSPARYGNLPVANTHRSAADVNTVYSDKIQTIPVVKAH 538
Qy 539 TLVSGTTVTKGPGTGNILKRTSSGPLAVTSVKSPLSQRYRARIYASTTNLRLFT 598
Db 536 MLVGGSVVQPGFTGCDILKRTNPSILGTFAVTNGSLSQRYRARIYASTTNLRLFT 598
Qy 599 ISGTIRYSINVKTNMGKDDLTNTFDLATIGTAPTFSYSDSLTGWGADSPASGGVYVD 658
Db 595 YLGDTTEKRNPNKMTNGASLTETTFKFSPIFDQFRETQDKILLSMGDFSSGGVYID 654
Qy 659 KPELIPVNAATFEABEDLDVAKAVNGLFTSKDALQTSVTDYQVQNAANLVECLSDLYP 718
Db 655 RIEPIVDETYEAEQLEAKKAVNALFTNFKGLRPGVTDYEVQNAANLVECLSDLYP 714
Qy 719 NEGRMLWDAVKEARLVQANLLODTCFNKINGENGWGTSGTIEVAGDVLFKDRSLRT 778
Db 715 NEKRLLFDAVREAKRLSGARNLQDPDFQREINGENGWAASTGIEIVEGDAVFKGRYLRLP 774
Qy 779 SAREIDTETPTLYQOIBESLAKPYTRYKLGPIGSSQDLEIKLIRHRANOIVKNVPDN 838
Db 775 GAREIDTETPTLYQOIBESLAKPYTRYKLGPIGSSQDLEIKLIRHRANOIVKNVPDN 834
Qy 839 LLPDVLVNSCGGIDRCSEQQYDANLALNNGENGNWSSDASHFPHIDTGEIDLMENT 898
Db 835 LLPDVLVNSCGGIDRCSEQQYDANLALNNGENGNWSSDASHFPHIDTGEIDLMENT 888
Qy 899 GIWVFKIPTNGYATLGNLELVEBGLSGETLERAQQOQQOQODKWKARKGASEKAYYA 958
Db 889 GIWVFKIPTNGYATLGNLELVEBGLSGETLERAQQOQQOQODKWKARKGASEKAYYA 948
Qy 959 AKQADRLPADYQDQKLNQVNSDMLAQNVLQSPITVYNDALPEIPGANYTSPTELTN 1018
Db 949 SKQAVELVADYQDQQLNPDVEITDLTAAQDLQSIPIVYVYNEFPEIPGANYTSPTELTD 1008
Qy 1019 RLQOANLYDLRNPAINPGRNGLSDWNATSDVNVQQLSDPTSVLVI PNNSQVSOQPTVQ 1078
Db 1009 RLQOANLYDLRNPAINPGRNGLSDWNATSDVNVQQLSDPTSVLVI PNNSQVSOQPTVQ 1068
Qy 1079 PNRYVLRVTRARKEGVDGYIIRDGANQTELTFTNI CDDDTGVLSADQTSYITKTVFT 1138
Db 1069 PNQRYVLRVTRARKEGVDGYIIRDGANQTELTFTNI CDDDTGVLSADQTSYITKTVFT 1128
Qy 1139 PSTEQWIDMSRTEGVFNIESVELVLEE 1167
Db 1129 PYTDQWIBISEGTFTYIESVELVLEE 1157
```

RESULT 5

```
US-07-915-203-2
; Sequence 2, Application US/07915203
; Patent No. 5359048
; GENERAL INFORMATION:
; APPLICANT: Ohba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Reiichi
; APPLICANT: Suzuki, No. 5359048ukazu
; APPLICANT: Ogiwara, Katsuooshi
; APPLICANT: Sakanaka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawasugi, Tadaaki
; TITLE OF INVENTION: No. 5359048el Microorganism and Insecticide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/915,203
; FILING DATE: 19920723
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/K 301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-915-203-2
```

Query Match 60.6%; Score 3662.5; DB 1; Length 1149;
Best Local Similarity 63.8%; Pred. No. 4.6e-260;
Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

```
Qy 1 MSPNQNEYIILDASSSTVSNDNSVRYPLANDQTTTLQNNYKYDLPMSGENSEBPCNP 60
Db 1 MSPNQNEYIILDASSSTVSNDNSVRYPLANDQTTTLQNNYKYDLPMSGENSEBPCNP 60
Qy 61 ETPISS-STVOTGIGIVGOVLGALVPPAGQASFYFIVGQLPSPSTVSVMIMKQV 119
Db 61 GTFISQDQAVGTGIDIVSTIISGLGIPVLGEVSILGSLGLLPSNNENWQIPMRVE 120
Qy 120 DLIDQKITDSVRKTALAGLQGLDLYQKSLKNLENRNDTRPARSVVTVQYIALELDF 179
Db 121 ELIDQKILDSVRSRAJADLANSVIAYVYQNALEDWKNPHSTRSAALVKERFPGNAEIL 180
Qy 180 VAKIPSPAIQORVPLLSVYAQAANLHLLLRDASIFGAEWGPTTGEISFYDRQVTRTA 239
Db 181 RTNMGSPSQNTYETPLLPYTAQAASLHLLVRDQVYIKGKNGYPQNDIDLYFKSQVSYTA 240
Qy 240 QYSDYCVKVTYGLDKLGTNAASLKYHOPRREMTLLVLDLVALFPNYDTRTYPIETTA 299
Db 241 RYSDHCQVMTNAGLNKLRGTAKQWVDYNNFRFRFRFRFRFRFRFRFRFRFRFRFRFR 300
Qy 300 QLATREVVYTDPIVFNRETSGGFCRWSLNSDISFSEVESAVIRSPHLPDIL 349
Db 301 ELTREIETDPV-----GSVTVQSSTLISWYDMIPALPSPSTLEN-LLRKDPFPFTLL 352
```


Qy 350 SEIEFYTTAGLPLNNT-EYLEYVGHSHIKYKNTNASSALERNYGTITSNKIKYVDLANK 408
Db 353 QEIRMYTS---FRQGTIEYNYNGGQRLTSLYIGSSP--NKYSGLVLAGAEDIIIPVGON 407
Qy 409 DIPQVRSIGADLANAYAQVYGPVYASFTLLDKNTGSGVGGFTYKPKHTTMOVCTQYNT 468
Db 408 DIYRV--VMTYIGRYTNSLLGNVPVPF-YFSNNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPPENEPLSRGYSRHSLSHTSYSPKSNASSPARYGNLVPFVWTHRSADVNTVY 525
Db 455 IDSGBELTYEN---YQSYSHRVSYITSPEIKSTGTV--LGWVPIFGWTHSSASRNPIY 509
Qy 526 SDKITQIPVKAHTLVSTTVIKGPB-PTGNNILKRTSSGPLAY---TSVSVKPLSORY 581
Db 510 ATKISQIPINKASRTSGGAVNMFQGLYNGGPMVKLSGSGSQVINLRVATDAKG-ASQRY 568
Qy 582 RARIYASTNNLRPLVNTS-----GTRIYSINVNKTWNGDDLTENTPDLATIGTA-P 633
Db 569 RIRIYASDRAGKT--TISRSRPNPATYSIASIAYTNTWNTNASLUTYSTFAYASGPNL 626
Qy 634 TFSNYSLSLTVGADSPAGSGVYVDKFLIPVNAATFEAEEDLDVAKKAVNGLFTSKDAL 693
Db 627 GISGSRFTDISITKEGAANLYIDRIEFIPVNTLFRAEEDLDVAKKAVNGLFTNERDAL 686
Qy 694 QTSVTDVQVQAAANLVECLSDDELIPNEKRLMDVAKEAKRLVQARNLLQDTGPNRINGEN 753
Db 687 QTSVTDVQVQAAANLVECLSDDELIPNEKRLMDVAKEAKRLVQARNLLQDTGPNRINGEN 746
Qy 754 GWTGSGTGLEVAEGDVLKFORSLRTSAREIDETPTVLYOOIHDESLAKPYTKLKGFI 813
Db 747 GWTGSGTGLEVAEGDVLKFORSLRTSAREIDETPTVLYOOIHDESLAKPYTKLKGFI 806
Qy 814 GSSQDLLEIKLRHANOIYKXVNPONLLPDVLPVNSCGGIDRCSEQQYVDANLALENNGEN 873
Db 807 GSSQDLLEIKLRHANOIYKXVNPONLLPDVLPVNSCGGIDRCSEQQYVDANLALENNGEN 866
Qy 874 GNMSSDSHAFPHIDTGBIDINENGTIWWPKIPTNGYATLGNLVELVEGFLSGETLER 933
Db 867 GNMSSDSHAFPHIDTGBIDINENGTIWWPKIPTNGYATLGNLVELVEGFLSGETLER 926
Qy 934 AQOGEQWQDQWARKRGAKEYAAKQAIIDRLPADYODOKLNSGVMSDMLAAQNLVQS 993
Db 927 AQOGEQWQDQWARKRGAKEYAAKQAIIDRLPADYODOKLNSGVMSDMLAAQNLVQS 986
Qy 994 IPYVNDALPEIPGNYTSTFELNRLQAAWNLXDLRNPAGDPFRNGLSDWATSDVNV 1053
Db 987 IPYVNDALPEIPGNYTSTFELNRLQAAWNLXDLRNPAGDPFRNGLSDWATSDVNV 1046
Qy 1054 QOLSSTSVLTPNNNSQVSOQFTVPNTRYVLRVTRARKEGVDGVIIRIDGANOTETLTF 1113
Db 1047 QOLSSTSVLTPNNNSQVSOQFTVPNTRYVLRVTRARKEGVDGVIIRIDGANOTETLTF 1106
Qy 1114 NICDDDTGVLSDQTSYITKTVEPTSPTEQWIDMSETTE 1152
Db 1107 NICDDDTGVLSDQTSYITKTVEPTSPTEQWIDMSETTE 1145

RESULT 6

US-08-272-887-2
; Sequence 2, Application US/08272887
; Patent No. 5747450
; GENERAL INFORMATION:
; APPLICANT: Ohba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Reichi
; APPLICANT: Suzuki, No. 5747450ukazu
; APPLICANT: Ogiwara, Katsutoshi
; APPLICANT: Sakanaoka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawasugi, Tadaaki
; TITLE OF INVENTION: No. 5747450el Microorganism and Insecticide

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/272.887
FILING DATE: 08-JUL-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/915,203
FILING DATE: 23-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: M/K 301
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1149 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-272-887-2

Query Match 60.6%; Score 3662.5; DB 1; Length 1149;

Best Local Similarity 63.8%; Pred. No. 4.6e-260; Indels 61; Gaps 21;
Matches 752; Conservative 109; Mismatches 257;
Qy 1 MSPNNQNEYELDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSBGNBELFGNP 60
Db 1 MSPNNQNEYELDASSSTSVSDNSVRYPLANDQTTTLQNNYKDYLRMSBGNBELFGNP 60
Qy 61 ETPFSS-STVGTGIGVQVLCALGVPPAGQIASFYSPVGLWESSITVVKWEMWKQVE 119
Db 61 GTFISAQDAVGTGDIVSTIIISGLGIPVLGVSILGLLWPSNNENVMQIFMNRVE 120
Qy 120 DLIDQKITDSVRKTAAGLQGLDGLDLYVQKSLKNLENRNDTRARSVVVTVYIALLELDF 179
Db 121 ELIDQKITDSVRSAIADLANSRVAYEQNALEDWRKNPHSTRSAALVKERPGNAREIL 180
Qy 180 VAKIPSPAIQSOEVPLLSVYAAANLHLLLRDASIFGAEWGFTPGIBISTFYDRQVTRTA 239
Db 181 RTNMGSPSQTYETPFLPTTAAQASLHLLVNRDVQIYKGEWGPQNDIDLFTYKBOVSXTA 240
Qy 240 QYSDYCVKWTYGLDKLGTNAASWLKHQFRRENTLLVLDLVALPNNYDTTYPIETTTA 299
Db 241 RYSDHCYQWYNAGLNKLGTCAGKQWVYNNRFRRENVMVLDLVALPNNYDARIYPLETNA 300
Qy 300 QLTREYVTDPIVFNRETSGGFCRRWSLN---SDI-----SFSEVESAVIRSPHLPDIL 349
Db 301 ELTREIFTDPV-----GSVYTCSSSTLSIWDMPAALPSFTLEN-LLRKDPDFTLL 352
Qy 350 SEIEFYTTAGLPLNNT-EYLEYVGHSHIKYKNTNASSALERNYGTITSNKIKYVDLANK 408
Db 353 QEIRMYTS---FRQGTIEYNYNGGQRLTSLYIGSSP--NKYSGLVLAGAEDIIIPVGON 407
Qy 409 DIPQVRSIGADLANAYAQVYGPVYASFTLLDKNTGSGVGGFTYKPKHTTMOVCTQYNT 468
Db 408 DIYRV--VMTYIGRYTNSLLGNVPVPF-YFSNNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPPENEPLSRGYSRHSLSHTSYSPKSNASSPARYGNLVPFVWTHRSADVNTVY 525

```

Db 455 IDSGEELTYEN---YQSYSHRVSVYITSPFEIKSTGGTV--LGVVPFGWTHSSASRNNPIY 509
Qy 526 SDKITQIPVKAHTLVSGTTVIKPGP-FTGNNILKRTSSGPLAY--TSVSVKSPLSQRY 581
Db 510 ATKISQIPINKASRTSGAVNMFQGLYNGGVMKLSGSGQVINLRVATDAKG-ASQRY 568
Qy 582 RARIRYASTTNLRFLFVTIS-----GTRIYSINVKNKMGDDLTFTPTPLATIGTA-P 633
Db 569 RIRIRYASDRAGKF--TISRSRSPENPATYSASIAFTWMTNSTNASLTYSTFAYSRGPINL 626
Qy 634 TFSNYSPLTVGADSPASGGVYVDKFLI PVNATFEAREDDVDVAKAVNGLPTSKDAL 693
Db 627 GISGSRFTDISITKEGAANLYIDRIEFPNTLFEAREDDVDVAKAVNGLPTNEKDAL 686
Qy 694 QTSVTDYQVQAANLVECLSDLEYPNKRMWDVAKKRLVQARNLLQDTGPNRINGEN 753
Db 687 QTSVTDYQVQAANLIECLSDLEYPNKRMWDVAKKRLVQARNLLQDTGFNRINGEN 746
Qy 754 GWTGSGTIEVAGDVLFKDRLSLRTSAREIDTETPTLYQOIHDESILKPYTRYKLGPI 813
Db 747 GWTGSGTIEVVEGDLFKDRLSLRTSAREIDTETPTLYQOIHDESILKPYTRYKLGPI 806
Qy 814 GSSODLEIKLRHRANQIVKRVNPNLLPDVLPVNSCGGIDRCSEQVYVDANLLENNGEN 873
Db 807 GSSODLEIKLRHRANQIVKRVNPNLLPDVLPVNSCGGIDRCSEQVYVDANLLENNGEN 866
Qy 874 GNMSSDSHAFSHDITGHEIDLNTNGIWWFKIPTTNGIYATLGNLELVEGFLSGETLER 933
Db 867 GNMSSDSHAFSHDITGHEIDLNTNGIWWFKIPTTNGIYATLGNLELVEGFLSGETLER 926
Qy 934 AQOQEQWQDMARKRGASEKAYTAQAIDRLFADYQDQKLGNGVMSDMLAAQNLVQS 993
Db 927 AQOQEQWQDMARKRAASEKAYTAQAIDRLFADYQDQKLGNGVMSDMLAAQNLVQS 986
Qy 994 IPYVYNDALPEIPGMYNTSPTELTNRLOQANWLDLRNAIPNGDFRNLSDMNATSDVNV 1053
Db 987 IPYVYNDALPEIPGMYNTSPTELTNRLOQANWLDLRNAIPNGDFRNLSDMNATSDVNV 1046
Qy 1054 QQLSDTSVLVLPNNSQVSQFTQPNRYVRLVTRARKEGVGDGVVIRIDGANQETILTF 1113
Db 1047 QQLSDTSVLVLPNNSQVSQFTQPNRYVRLVTRARKEGVGDGVVIRIDGANQETILTF 1106
Qy 1114 NICDDTGVLSADQTSYITKTVEPTSPTEQVWIDMSETE 1152
Db 1107 NICDDTGVLSADQTSYITKTVEPTSPTEQVWIDMSETE 1145

```

RESULT 7

```

US-08-789-449-2
; Sequence 2, Application US/08789449
; Patent No. 5824878
; GENERAL INFORMATION:
; APPLICANT: Obba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Relichi
; APPLICANT: Suzuki, No. 5824878ukazu
; APPLICANT: Ogiwara, Kazunobu
; APPLICANT: Sakanaka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawaasugi, Tadaaki
; TITLE OF INVENTION: No. 5824878el Microorganism and Insecticide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/789,449
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,887
; FILING DATE: 08-JUL-1994
; APPLICATION NUMBER: US 07/915,203
; FILING DATE: 23-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/K 301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-789-449-2

Query Match 60.6%; Score 3662.5; DB 2; Length 1149;
Best Local Similarity 63.8%; Pred. No. 4.6e-260;
Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

Qy 1 MSPNNQNEYILDASSSTVSVDNSVRYPLANDQTTTLQNNMYKDYLRMSGENPFLPGNP 60
Db 1 MSPNNQNEYILDASSSTVSVDNSVRYPLANDQTTTLQNNMYKDYLRMSGENPFLPGNP 60
Qy 61 EFTISS-STVGTGIGVGOVLGALGVPPAGQIASFYPIVGQLPFSSTVSVMEMIKQVE 119
Db 61 GTFISAQDAVGTDIVSTIISGLGIPVLGEVPSILGSLGLANTNNENVMQIPANRV 120
Qy 120 DLIDQKITDSVRKTALAGLQGLDGLVYQSKLKNLNRNDTRARSVVVTOYIALELDF 179
Db 121 ELIDQKILDSVRSRAIADLANSRIAVEYVQNALEDNRKNPHSTRSAALKRFGNAEAIL 180
Qy 180 VAKIPSPALISQGVPLLSVYQAANLHLLLRDASIFGAEWGFTFCEISTEYDROVTRTA 239
Db 181 RTNMGSPSQTYETPLLEFYAQAASLHLLVWRDVOIQYKMGWGPQNDIDLFKEQVSYTA 240
Qy 240 QYSDYCVKWNVTGLDKLGTNAASWLKYHOFRENTLLVLDLVALFPNTYDTRTYIETTA 299
Db 241 RYSDHCWQVYNAGLNKLGTCAGKQWVDNRRFERNNVWLDLVALFPNTYDARIYPLETNA 300
Qy 300 QLTREVYTDPIVFNRETSGGFCRRWSLN-----SDI-----SFSEVESAVIRSHULFDIL 349
Db 301 ELTREIFTDPV-----GSYVTGQSSTLISWYDMIPALPSPSTLEN-LLRKDPDFTL 352
Qy 350 SEIEFTYTRAGLPLNNT-EYLEVWVGHISIKYKNTNASSALERNYGTITSNKIKYYDLANK 408
Db 353 QSEIRMYTS---FRQNGTIEYNYNGGQRLTSLYIYGSSP--NKYSGVLAGEADIIIPVQGN 407
Qy 409 DIFQVRSILGADLANYYAOVYGVYASFTLLDRNTGSGSVGGFTYKPKHTTMOVCTQNYNT 468
Db 408 DIYRV-VWTVIGRTNSLLGVNPVTF-YFSNNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPPENEPLSRGSHRLSHITSYTSFSKQNASPPARYGNLPVPAWTHRSADVNTVY 525
Db 455 IDSGBELTYEN---YQSYSHRVSVYITSPFEIKSTGGTV--LGVVPFGWTHSSASRNNPIY 509
Qy 526 SDKITQIPVKAHTLVSGTTVIKPGP-FTGNNILKRTSSGPLAY--TSVSVKSPLSQRY 581
Db 510 ATKISQIPINKASRTSGAVNMFQGLYNGGVMKLSGSGQVINLRVATDAKG-ASQRY 568
Qy 582 RARIRYASTTNLRFLFVTIS-----GTRIYSINVKNKMGDDLTFTPTPLATIGTA-P 633
Db 569 RIRIRYASDRAGKF--TISRSRSPENPATYSASIAFTWMTNSTNASLTYSTFAYSRGPINL 626

```

QY 634 TFSNYSLSLTVGADSFASGEVYVVDKFLIPVATFEABEDLDVAKAVNGLFTSKDAL 693
DB 627 GIGSSRTFDSITKEAGANLYIDRIEFIPVNTLFEABEDLDVAKAVNGLFTNEKDAL 686
QY 694 QTSVTDYOVQOANLVECLSDLYPNKRLMDVAKKRLVQARNLLQDTGPNRINEN 753
DB 687 QTSVTDYOVQOANLVECLSDLYPNKRLMDVAKKRLVQARNLLQDTGPNRINEN 746
QY 754 GWTGSGTGEVAGDVLKDRSLRLTSAREIDTETPTLYQQIDSLKPYRYKLKGF 813
DB 747 GWTGSGTGEVAGDVLKDRSLRLTSAREIDTETPTLYQQIDSLKPYRYKLKGF 806
QY 814 GSSQDLEIKLIRHRANQIVKVPDNLDPVLPVNSCGGIDRCSEQOYVDANLLENNGEN 873
DB 807 GSSQDLEIKLIRHRANQIVKVPDNLDPVLPVNSCGGIDRCSEQOYVDANLLENNGEN 866
QY 874 GMSDSDSHAFPHIDTGBIDNENGIWVPKIPITNGVATLGNLEVBEGPLSGETLER 933
DB 867 GMSDSDSHAFPHIDTGBIDNENGIWVPKIPITNGVATLGNLEVBEGPLSGETLER 926
QY 934 AQOEOQWQDQWARKGASERKAYAAKQAI DELFADYQDQKLNKSGVEMSDMLAQNLYQS 993
DB 927 AQOEOQWQDQWARKGASERKAYAAKQAI DELFADYQDQKLNKSGVEMSDMLAQNLYQS 986
QY 994 IPYVNDALPEIPGNNTSFTLNLRLQOANLVDLRNAINPNDPRNGLSNWNATSDVNV 1053
DB 987 IPYVNDALPEIPGNNTSFTLNLRLQOANLVDLRNAINPNDPRNGLSNWNATSDVNV 1046
QY 1054 QOQSDTSVLVLPNWNQSVQOFTVPYRYVLRVTRARKEGVGDGVVIRDCANQOTETLTP 1113
DB 1047 QOQSDTSVLVLPNWNQSVQOFTVPYRYVLRVTRARKEGVGDGVVIRDCANQOTETLTP 1106
QY 1114 NICDDTGVLSADQTSYTKTVEFTPTSTEQWIDMSET 1152
DB 1107 NICDDTGVLSADQTSYTKTVEFTPTSTEQWIDMSET 1145

RESULT 8

US-08-315-468-4
; Sequence 4, Application US/08315468
; Patent No. 5554534
; GENERAL INFORMATION:
; APPLICANT: Michaels, Tracy Ellis
; APPLICANT: Foncecra, Luis
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Process for Controlling Scarab Pests
; TITLE OF INVENTION: with Bacillus thuringiensis Isolates
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL USA
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/315,468
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,941
; FILING DATE: 01 FEB 1993
; APPLICATION NUMBER: 07/828,430
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/808,316
; FILING DATE: 16-DEC-1991

ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA73.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1169 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: 50C
; IMMEDIATE SOURCE:
; LIBRARY: LambdaGEM-11 library of L. Foncecra
; CLONE: 50C(b)
US-08-315-468-4

Query Match 52.2%; Score 3152; DB 1; Length 1169;
Best Local Similarity 54.0%; Pred. No. 1.4e-222;
Matches 653; Conservative 179; Mismatches 294; Indels 84; Gaps 25;

QY 1 MSPNNQNEYILDDASSSTVSVDNSVRYPLANDOTTTLQNMNMYKDYLRMSSEGENPELPGNP 60
DB 1 MSPNNQNEYILDDASSSTVSVDNSVRYPLANDOTTTLQNMNMYKDYLRMSSEGENPELPGNP 60
QY 61 EFTFIS-SSTVQIGIVGQVLGALGVPFAGQIASFYISFVIGQMLPSSSTVSVMIMHMQVE 119
DB 61 EFTFIS-SSTVQIGIVGQVLGALGVPFAGQIASFYISFVIGQMLPSSSTVSVMIMHMQVE 120
QY 120 DLIDOKITDSVRKALAGLOGLDGLDYQSKLKNWLENDRTRARSVVVTOYIALELDF 179
DB 121 ELINOKIAEYARNKALSELEGNYYQLYLTALBEWKENPNNGSRALDRVNRFEILDSLF 180
QY 180 VAKIPSPAIQOEVPFLSVYQAANLHLLLDASIFGAEMGFTPGIEISTFYDROVTRTA 239
DB 181 TOYMESFRVTFEVPFLVYVYTAANLHLLLDASIFGAEMGFTPGIEISTFYDROVTRTA 240
QY 240 QYSDYCVKMYNTGLDKLGTNAASVLYKHQFEREMTLVLDLVALFPNYDTPTPIETTA 299
DB 241 EYSDHCVKMYETGLAKGSSAKQWIDYNQFEREMTLVLDLVALFPNYDTPTPIETTA 300
QY 300 QLTREVTYDPIVFNRETSGGFCRRWLSNDSISFSVESAVIRSPHLPDLSIEPVTTRA 359
DB 301 QLTREVTYDPIVFNRETSGGFCRRWLSNDSISFSVESAVIRSPHLPDLSIEPVTTRA 356
QY 360 GLPLNTEYLVVWCHSIKYKNTNASSALEENYGT-----ITSNKIKYDLANKDIQVFR 414
DB 357 S--FTSDRYMYWAGHQSISYKHIGTSSTFTQYGNQNLQSTN-----FDFNYDIYKTL 410
QY 415 SIGADLAN-----YYAQVGVPIASFTLLDKNMGSGVGGFTYKSPHTTMQVCTQNTYI 469
DB 411 SNGAVLLDIVPGYTYTFPGMPETEFPMVNLNNTRKT--LTY-KP-ASKDIIDRTDSE 466
QY 470 DEIPPE--NEPLSRGYSHRLSHITSYSPKNAASPARGNLVPANTHRSADVNTVYSD 527
DB 467 LELPPEPSTGQPNYESYHRLGHIT--FYSSSTST-----YVPVFWTHRSADLTNTVKS 520
QY 528 KITQIPVVKATLVSGTTVKGPGTGNILKTSSTGGLAYTSVSVKSPLSQRYARIRY 587
DB 521 EITQIPGKSGSTIGRNTIIRGRTGGDLVALTR--IGSCFQMPFESQRIIRY 578
QY 588 AS--TTNRLRF-VTISGTRIYSINVKNT-MNKGD-DLTPNTPLDLAGTFTFSYSDSL 642
DB 579 ASNETSYISLYGLNQSGT-----LKPNTYSNKNENDLYND-----PKYIEYPRVI 625

QY 945 MARKGASAKYAAKQADLFDADYODOKLNGVMSDMLAQNLOSIPVYNDALPE 1004
Db 935 LGRKRAIRIDRVYLAQAQNLHLPVYDQOQNLNPEIGLAELNEASNLVESISGVYSDTLQ 994
QY 1005 IFCGNTSTFELNRLQQAQNLNLDLNAIPNGDFRNGLSDNWATSQVNVQQLSDTSVLVI 1064
Db 995 IFCINVEIYTELSDRLQQAQSYLYTSRNAVQNGDFNSCLDSWNTTMDASVQDCGNMFLVL 1054
QY 1065 PNWNSQVSOQTPVQPHRYVLRVYARKEGVDGYVIRIDGANTETILTNNICDDDTGVL 1124
Db 1055 SHWDAQVSOQLRVNPNCKYVLRVYARKEGVDGYVIRIDGANTETILTNNICDDDTGVL 1124
QY 1125 ADQTSYITKTVPRTPTFEQWIDMSTEGVFNESVELVLEER 1167
Db 1115 VNDNSYITTEVVPYPTKMMWVSESEGSFYIDSIEPIETQE 1157

RESULT 11

US-08-455-838-5
; Sequence 5, Application US/08455838
; Patent No. 6028246
; GENERAL INFORMATION:
; APPLICANT: Lambert Bart
; APPLICANT: Jansens, Stefan
; APPLICANT: Van Audenhove, Katrien
; APPLICANT: Peferoen, Mathix
; TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND
; TITLE OF INVENTION: THEIR INSECTICIDAL PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,838
; FILING DATE: 31-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/01820
; FILING DATE: 12-JULY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93400949.9
; FILING DATE: 09-APR-1993
; APPLICATION NUMBER: EP 92402358.8
; FILING DATE: 27-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 2121-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-838-5

Query Match 51.7%; Score 3123; DB 3; Length 1157;
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

QY 1 MSPNNONEVEILDASSSTSVSDNSVRYPFLANDOTTTLQNNMYKYDLRMSGEGNPELFCNP 60
Db 1 MNRNNQNEVEIILDAHPCCPSDDRYPLASDPNAAALQNNMYKYDLQNTDEYDTSYNP 60
QY 61 ETPISS-STVQTGIGIVGVVAGALGVPFAGQIASFYSPFVQOLMSPSSVTSVWEMIMQV 119
Db 61 SLISGRDAVQVATVVGRIILGALGVFPFSGOIVSFYQFLNLTLPVNDTATWEAFMRQVE 120
QY 120 DLIIDOKITDSVRKTALAGLQGLGVYQKSLKWLLENNDTARSVVVYTOYTALELDP 179
Db 121 ELVNOQITFEARNOALRLQGLGDSFNVYQSLQNLWADNRDNLNLSVRAQFIALDLDP 180
QY 180 VAKIPSAISGOEVPILLSVYAQAANLHLLLRDASIFCAEWGPTPGHISTFYDQVTRTA 239
Db 181 VNAIPFAVNGQVPLLSVYAQVAVNLHLLLRDASLFGEGWPTQGHISTFYDQVLETA 240
QY 240 QYSDYCVKWNVTGLDKLKGNTAASWLKYHOFREEMTLVLVLDLVALFPNYOTRTPIETTA 299
Db 241 KYTNYCETWYNTGLDLRGNTESMLRYHQFRREMTLVLDVVALFPYDVLRYTGSNP 300
QY 300 QLTREVTYDPIVFNRETSGGFCRRNSLNSDISFSEVESAVIRSPHLFDILSEIFYTTTRA 359
Db 301 QLTREVTYDPIVFNRETSGGFCRRNSLNSDISFSEVESAVIRSPHLFDILSEIFYTTTRA 359
QY 360 GLPLNTEYLEVWVGHSHIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFVRS 416
Db 360 -FPV-SSNFMVYWSGHTLRRSYLNDSSAVQSDSLGTLITTRATINPGVDGTR-----IBST 413
QY 417 GADLANVYQVYGVYASFTL--LDKNTGSGVGGFTYKPHHTMQVCTQNYNTIDEIPP 474
Db 414 AVDFRSALIGIYGVNRASFPVPGGLFNGTTPSPANGG-----CRDLYDNDLPP 461
QY 475 ENEPLRGYS-HRLSHITSYSPSKN-ASSPARVGNLPVPAWTHRSADVTNIVYSDKKTQI 532
Db 462 DB---STGSSHRLSHVTFPSPQTNQAGSIANAGSVPTVYVTRDVLNLTITFNRI 518
QY 533 FVVKATLVSGTIVIKGPGTGNILKRTSSGPLAYTSVSVKSPLSORYARIRIYASTN 592
Db 519 PLVKASAPVSGTIVIKGPGTGGILRRITNGTFTGLRTVTVNSPLTQOYRLRVFAS 578
QY 593 LRLFTVITSGTRIYSINVNKTMKGGDLTFNTP---DLATIG---TATFSGNYSLSLTGA 646
Db 579 FSIRVLRGVSGIGDVLRLGSTMNGOELTYESPFTREPTTTPGPNPPTFTQAOEILTV 638
QY 647 DSPASGEVTVDFKPLIPVNAFTEAEEDLDVAKAVNGLTFSKDALQTSVTDVQVNOAA 706
Db 639 EGVSTGGEYVIDRIBIVPVNPAEAEEDLEAKKAVASLFTRTDGLQVNVTDYQVDOAA 698
QY 707 NLVECLSDLELYPNEKRLMDAVKEAKLVQARNLQDQTFGRNNG---ENGWGTSGTGEVA 764
Db 699 NLVSCLSDBEQYGHDKKMLLEAVRAAKRLSRERMLLQDPDTINSTENGWKAENGVTIS 758
QY 765 EGDVLFDRSLRLTSAREIDTETPTLYQQIDESLLKPYTRYKLGKFIGSSQDLKLI 824
Db 759 EGGPFKGRALQASAR---ENVPTIYQKVDAVLKPYTRYLDCGVKSSQDLKLI 814
QY 825 RHRAQIVKXVNDLPLPDVLPVNSCGGIDRCSCQVYDANLALNNGENGWMSDSDSHAPS 884
Db 815 HHKVLHVLKXVNDLPLPDVLPVNSCGGIDRCSCQVYDANLALNNGENGWMSDSDSHAPS 874
QY 885 FHLDTGEIDLNENTGIWVVKIPPTNGYATLGNLELVEEGPLSGETLERAQOQOQOQ 944
Db 875 SYINTGLNNSVDQGIWVVKVTTDGYATLGNLELVEEGPLSGESLERQORDAKKNAB 934
QY 945 MARKGASAKYAAKQADLFDADYODOKLNGVMSDMLAQNLOSIPVYNDALPE 1004
Db 935 LGRKRAIRIDRVYLAQAQNLHLPVYDQOQNLNPEIGLAELNEASNLVESISGVYSDTLQ 994
QY 1005 IFCGNTSTFELNRLQQAQNLNLDLNAIPNGDFRNGLSDNWATSQVNVQQLSDTSVLVI 1064
Db 995 IFCINVEIYTELSDRLQQAQSYLYTSRNAVQNGDFNSCLDSWNTTMDASVQDCGNMFLVL 1054
QY 1065 PNWNSQVSOQTPVQPHRYVLRVYARKEGVDGYVIRIDGANTETILTNNICDDDTGVL 1124
Db 1055 SHWDAQVSOQLRVNPNCKYVLRVYARKEGVDGYVIRIDGANTETILTNNICDDDTGVL 1124
QY 1125 ADQTSYITKTVPRTPTFEQWIDMSTEGVFNESVELVLEER 1167
Db 1115 VNDNSYITTEVVPYPTKMMWVSESEGSFYIDSIEPIETQE 1157

Db 1055 SHMDAQVQQLEVNPNCKYVLRVARKVGGDGVYTRDGAHQHQBTLTFRACDYVNGTY 1114
Qy 1125 ADQSYTKTKVEPTSPTEQVWIDMSSETGVNIESVELVLEE 1167
Db 1115 VNDNSYITEEVVFPETKHMVSESGSEGYIDSIETQOE 1157

RESULT 12
US-09-019-809-5
; Sequence 5, Application US/09019809
; Patent No. 6143550
; GENERAL INFORMATION:
; APPLICANT: LAMBERT, BART
; APPLICANT: JANSSENS, STEFAN
; APPLICANT: VAN AUDENHOVE, KATRIEN
; APPLICANT: PEPEEROEN, MARINX
; APPLICANT: VAN RIE, JEROEN
; APPLICANT: VAN AARSSEN, ROEL
; TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND THEIR
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,809
; FILING DATE: 02-FEB-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 2121-135P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-019-809-5

Query Match 51.7%; Score 3123; DB 3; Length 1157;
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

Qy 1 MSPNNQNEYITLDASSSTVSVDNSVRYPVPLANDQTTLLQNNMYKDYLRMSGEGNPELFCNP 60
Db 1 MNRNNQNEYITLDAPHCPCPSDDDDVRYPLASDPDPAALQNNMYKDYLTQMTDEDYDYSINP 60
Qy 61 ETFISS-STVQTGIGIVGQVLTGALGVPPAGQIASPYSFIVGQLMPSSTVSVMIMKQVE 119
Db 61 SLSISGRDAVQVLTAVGRILGALGVPPSGQIVSFYQFLNTLMPVNDTALWEAFMRQVE 120
Qy 120 DLIQKTDTSVRKTKALAGLOGDGLGVQVQSLKQNLNRNDTRARSVVVYQYALELDF 179
Db 121 ELVNNQITEFARNQALAFLOGDGSFNVYQSLQNLNLADRNDTRNLVSRYAQFIALDLDF 180
Qy 180 VAKTPSPAISGOEYPLLSVYAAQANLHLLLRDASIFGABWGFTPGESTPYDQVTRTA 239
Db 181 VNAIPFAVNGQQVPLLSVYAAQANLHLLLRDASLFGEGWGFTPGESTPYDQVTRTA 240

Qy 240 QYSDYCVKYNNTGLDKLGTNAASWLKHQPRREMTLLVLVLVALFPNYDTRTYPIETTA 299
Db 241 KYTNYCETWNTGLDRGNTGNTESWLRYHQPRREMTLLVLVLVALFPYDVRVLYTGSNP 300
Qy 300 QLTREVVTDPIVRENETSGGFCREWSANSIDISFSEVSASVIRSHPLFDILSIEPIETRA 359
Db 301 QLTREVVTDPIVFPNPPANVGLCRWGTNPNYNTFSELENAPIRPPLPRLNLSLTSSNR- 359
Qy 360 GLPLANTLEYLWYVGHSHIKYKNTNTASSALENRYGTITSNKIKY---YDLANKDIQVRSIL 416
Db 360 -PFV-SSNPDYWSGHTLRSLYNSLDAVQEDSYGLITTRATINPGVDGTR-----IEST 413
Qy 417 GADLANYAQQVYGPYASFTL--LDKNTGSGVSGFTYSKPHHTMQVCTQYNTIDEIPPP 474
Db 414 AVDFRSALIGYGVNRASFPVGGFLNGTISPPANGG-----CRDLYDNDDELPP 461
Qy 475 ENBPLSRGYS-HRLSHITSYSFSKN-ASSPARYGNLVPFAMTHRSADVTNTVYSKIQI 532
Db 462 DE---STGSSSTRLSHVTFPSQTNQAGSIANAGSVPTVWTRDVLNLTITPNTITQL 518
Qy 533 PVKHAHTLVSGTIVIKGPGFTGGNILKRTSSGPLYTSVSKSPLSQRYRARIYASTTN 592
Db 519 PLVKASAPVSGTIVLKGPGFTGGILRRITNGTFTGLRVTYNSPLTQYRLRVRPASTGN 578
Qy 593 LRLPVTISGTRIYSINVNKTKNGDDLTFTNTF---DLATIG---TAPFYSNYSDSLTVGA 646
Db 579 PSIRVLRGVSGIDVRLGSTMNRQELTYSEFTTTPGFPNPPFTTQOQELTVNA 638
Qy 647 DSPASGGEVYVDKELIPVNATFEAREDDLVAKVANGFLTSSKDALCTSVTDYOVNOAA 706
Db 639 EGVSTGGYIYDRIEIVFVNPAREAREDBAARKLRSERNLLQDPDFNTINSTENGKASNGVTIS 758
Qy 707 NLVECLSDLEYNPKRMMLDAVKEAKRLVQARNLLQDTGFNRING--ENGWGTSGTIEVA 764
Db 699 NLVSCLSDEQYGHDKMLLEAVRAAKRLSERNLLQDPDFNTINSTENGKASNGVTIS 758
Qy 765 EGVVLFKDRSLRLTSAREIDTETPTLYYQOIDSLLKPYTRYKLGKPGISSQDLEIKLI 824
Db 759 EGGPFFKGRALQLASAR----ENPTVIYQVDASVLKPYTRYLDGFPVKSQDLEIDL 814
Qy 825 RHRAQIVKVPDNLDPVLPVNSCGGIDRCSEQYVDANLALNNGNGNMSDSSHAPS 884
Db 815 HHKHLVKNVPDNLVSDTYSDCSGSLNRCDEQHVDMQLDAEHHPMDCEAAQTHEPS 874
Qy 885 PHIDTGEIDLMENTGIWVPKIPITTINGVATLGNLELVEEGLSGSETLERRAQOQOOWDK 944
Db 875 SYINTGDLNASVDQGIWVLKVRTTDDGYATLGNLELVEEGLSGSETLERRAQOQOOWDK 934
Qy 945 MARIKGASEKAYAAKQAIIDRLPADYQDOKLNSGVNSDMLAAQHLVQSIPIVYNDALPE 1004
Db 935 LGKRAEIDRVYLAQAQAINELPVDYQDQQLNPRIGLAEINEASNLVESISGVYSDTLQ 994
Qy 1005 IPGKNYTSFTBLTNRLQAQANLYDLRNIAPNGDFFNGLSDWNATSDVNVQQLSDTSVLVI 1064
Db 995 IPGNIYEITSELRLOQAQSYLTYSRAVQNGDFNSGLSDSWNTTADASVQDQGNHFLVL 1054
Qy 1065 PNNNSQVSSQQTQVQFNRYVLRVTRARKEGVGDGVYIIRDGANQTELTFTNICDDDTGVLS 1124
Db 1055 SHMDAQVQQLEVNPNCKYVLRVARKVGGDGVYTRDGAHQHQBTLTFRACDYVNGTY 1114
Qy 1125 ADQSYTKTKVEPTSPTEQVWIDMSSETGVNIESVELVLEE 1167
Db 1115 VNDNSYITEEVVFPETKHMVSESGSEGYIDSIETQOE 1157

RESULT 13
US-09-471-177-5
; Sequence 5, Application US/09471177
; Patent No. 6448226
; GENERAL INFORMATION:
; APPLICANT: LAMBERT, BART
; APPLICANT: JANSSENS, STEFAN

APPLICANT: VAN AUDENHOVE, KATHRIEN
APPLICANT: PEEREN, MARNIX
APPLICANT: VAN RIE, JEROEN
APPLICANT: VAN AARSSEN, ROEL
TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND THEIR
TITLE OF INVENTION: INSECTICIDAL PROTEINS.
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/471,177
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/019,809
FILING DATE: 02-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 2121-135P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-471-177-5

Query Match 51.7%; Score 3123; DB 4; Length 1157;
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

QY 1 MSPNNEYILASSSTVSNSVRYPLANDQTTTLQNNYKDYLSKSGRPELFGNP 60
DB 1 MNPNNNEYEIIDAPHGCPSSDDVRYPLASDPNAAALQNNYKDYLTQMTDEDYDTSYNP 60
QY 61 ETPISS-STVQTGIGVIGVGLGALGVPFAGQIASFYSFVQGLWPSTSVSWEMMKQVE 119
DB 61 SLSISGRDAVOTALTVVGRILGALGVPFSGQIVSFYQFLTLNLPVNDTAIMEAFMEQVE 120
QY 120 DLIDQKITSVRKTAGLAGLGLDGLDVYOKSLKNWLENRNDTRARSVVVTVYALELDF 179
DB 121 ELVNOQITEPARNOALRQLGDSFNVYQSLQNLADRNDTRNLNVRAQFIALDLDF 180
QY 180 VAKIPSPAGISQGVPLSVYQAQANLHLRLDASIFGAEGWGPCEISTFYDROVTRTA 239
DB 181 VNAIPLPFVANGQQVPLSVYAQAANLHLRLDASIFGAEGWGPCEISTFYDROVTRTA 240
QY 240 QYSDYCVKNTGLDKLKGTAASWLYKHOFREMTLLDLVALFPNYDTRYPIETTA 299
DB 241 KTYNCETWNTGLDRGNTGNTSWLYKHOFREMTLLDLVALFPNYDTRYPIETTSNP 300
QY 300 QLTREVVYTDPIVFNRETSGGFCRRWSLNSDISFSEVASVIRSHPLDILSIEFYETRA 359
DB 301 QLTREVVYTDPIVFNPPANVGLCRWGTNPYNTFSELENAFIRPHPLDRLNSLTISNR- 359
QY 360 GLPLNNTLEYLYVWGHISKYKNTNASSALERNYTTISNKIKY---YDLANKDIFQVRS 416
DB 360 -FPV-SSNFMVDYNSGHTLRSLYNSDASVQEDSYGLTITTRATINPGVDGTR-----IEST 413

QY 417 GADLANYAQQVGVYASFTL--LDKNTGSGSVGGFTYISKPHYTMQVCTQNTYNTIDISPP 474
DB 414 AVDFRSALIGIYGNRASFPVGGFLFNGTTPANGG-----CRDLYDTNDELPL 461
QY 475 ENEPLSRGYS-HLSHITSYSPSKN-ASSPARVGNLPEVPAWTHRSADVNTVYSDKLTQI 532
DB 462 DE---STGSSTHLSHVTFPSFOTNQAGSIANAGSVFTYVTRDVLNNTITFNRTIQL 518
QY 533 PVVKAHTLVSGTIVKPGFTGNILKRTSSGFLAYTSVSVKSPLSQRYRARIYASTTN 592
DB 519 PLVKASAPVSGTIVLAGPGFTGGILLRRTNGFTGLRVTVNSPLTQYQLRVRFASVGN 578
QY 593 LRLFTVTSIGRIYSINWNTMNGDDLTFNP---DLATIG---TATFSPNSLSLTGA 646
DB 579 FSIRVLRGVSGISGDVRLGSTMNRGQELTYBSPFTRBFTTGPFPNPPFTTQAQILTVNA 638
QY 647 DSPASGGEVYVDKFEIIPVNATPEAREEDLDVAKVANGLTFSKDALQTSVTDYVQNOAA 706
DB 639 EGVSTGGEYIIDRIEIVPVNPARABEDLEAKKAVASLFRTRDGLQVNVTDYVQNOAA 698
QY 707 NLVECLSDELYPNKEKRLMDAVKAKELVQARNLQDTGFNRING--ENGWGTSGTGEVA 764
DB 699 NLVSCLSDRQYGHDKMLLEAVRAAKRLSRERNLQDPDFTNTINSTEENGKASNGVTIS 758
QY 765 EGDVLPKORSRLTSAREIDTETPTVLYQOIDEBSLLKPTRYKLKGFISGSQDLEIKLI 824
DB 759 EGGPFPKRALQLASAR---ENYPTVYQKVDASVLKPTRYRLDGFVKSSQDLEIDL 814
QY 825 RHKANOIVKNVDPNLLPDVLPVNSCGSIDRCSQVQVYDANLALNNGENGNNMSDSDSHAPS 884
DB 815 HHKHLVKNVDPNLLVSDTYSDCSGINRCDSHQVDMQLDAEHHPMDCCERAAQTHERFS 874
QY 885 FHIDTGBIDLNENTGIWVVFKEIPTTNGYATLGNLELVEEGPLSGETTLERAAQOQWQDK 944
DB 875 SYINTGDLASVDQGLVWLKVTTDGYATLGNLELVEVGLSGESLEREQDRAKNAE 934
QY 945 MARKGASERKAYAAKQADRLPADYQDQKLSNGSVEMSLAAQNLVQSIPTVYNDALPE 1004
DB 935 LGRKRAIDRVYLAQAQAINHLFVDYQDQQLNPEIGLAEINAEASNLVESISGVYSDTLQ 994
QY 1005 IFCNNTSFTLNRLOQANWLYDLNNAIENGPRNGLSDWNATSDVNVQQLSDTSVLVI 1064
DB 995 IFCNNTSFTLNRLOQANWLYDLNNAIENGPRNGLSDWNATSDVNVQQLSDTSVLVI 1054
QY 1065 PWNNSVSOQFTVQPNRYVLRVTRKGVGDGVIIIRDGANQOTETLTFTNCDDDTGVL 1124
DB 1055 SHWDAQVSQQLRVNPNCKYVLRVTRKGVGDGVIIIRDGANQOTETLTFTNCDDDTGVL 1114
QY 1125 ADQTSYITKTVFTPTPTEQVWIDMSETEGVFNIESVELVEE 1167
DB 1115 VNDNSYITEBWFVFPETKHMWVEVSESGSFYIDSIEFIETQE 1157

RESULT 14

US-09-002-285-72
Sequence 72, Application US/09002285
Patent No. 6369213

GENERAL INFORMATION:

APPLICANT: Schnepf, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth B.
APPLICANT: Walz, Michelle
APPLICANT: Stockhoff, Brian
APPLICANT: Muller-Cohn, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA

ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/002,285
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 1156 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-002-285-72

Query Match 51.3%; Score 3100.5; DB 4; Length 1156;
Best Local Similarity 52.0%; Pred. No. 8.5e-219;
Matches 615; Conservative 196; Mismatches 330; Indels 41; Gaps 13;

QY 1 MSPNNQYEILDASSSTSVSDNSVRYPLANDQTTTLQWNYKDYLRMSGENPELFGNP 60
DB 1 MNRNNQYEIIDPHCGPCSDDDVRYPLASDPNAAALQNNYKDYLRMSGENPELFGNP 60
QY 61 EFTISS-STVOTGIGVQVGLGALGVPPAGQIASPVSFVGLQWPSSTVSVWEMIMKQVE 119
DB 61 SLISIGRAVQTLTVWGERILGALGVPPGQIVSPYQFLNTLWPNVDTALWEAPRQVE 120
QY 120 DLIDQKITSVRKTAGLQGLGDLGVYQSKLNLENRNDTRARSVVVQYIALLELDF 179
DB 121 ELVNCQITEPARNQALRQGLGDSFVYQSLQNLADNRNLSVVRQFIALDLDF 180
QY 180 VAKIPSPASQGVPLLSVYQAANLHLLLEDASIFGAENGFTPGESTFYDRQVYRTA 239
DB 181 VNAIPFVNGQVPLLSVYQAANLHLLLEDASIFGAENGFTPGESTFYDRQVYRTA 240
QY 240 QYSDCVKMYNTGLKAGTAAASMLKYHOFREMTLLVLDLVALFPNVDYRTYPIETTA 299
DB 241 KTYNCETWYNTGLDRKGTWESLWYHOFREMTLLVLDLVALFPNVDYRTYPIETTA 300
QY 300 QLTREVTYDPIVFNRETSGGFCRWSLNSDISFSVESAVIRSPHLPDLSIEFYTTTRA 359
DB 301 QLTREVTYDPIVFNRETSGGFCRWSLNSDISFSVESAVIRSPHLPDLSIEFYTTTRA 359
QY 360 GLPLNNTYLEWVWCHSIKYKYNASSALERNVYGTITSNKIKY---YDLANKDIFQVRS 416
DB 360 -FPV-SSNFMVWSGHTLRSYLNDSAVQEDSYGLITTTTRATINPGVDGNTN-IBST 413
QY 417 GADLANVYQVGVVPSFTL--LQNTGSGVSGGFTYSKPHYTMQVCTQNTYNTIDEIPP 474
DB 414 AVDFRSALIGYGVNRASFPVGLFNGTSPANGG-----CEDLYDNDLELPP 461
QY 475 ENEPLSRGVS-HRLSHITSYSPSKN-ASSPARYGNLPVFAWTHRSADVNTVYSDKITQI 532
DB 462 DE---STGSSHRLSHVTFPSQTQOAGSIANAGSVPTVYVTRRDVLDLNNITPNTIQL 518
QY 533 PVVKAHTLVSGTVIKGPGFTGCGNLIKRTSSGFLAYTSVSKSPLSQRYRARIYASTTN 592

DB 519 PLVXASAPVSGTTLKGGFTGGILRTTNGTFTGLRTVTVNSPLTORVRVRFPASSGN 578
QY 593 LRLFTVTSITRIYSINNVKTKGDDLTFTNTFDLATIGT-----AFTFGNYSDSLTVGAD 647
DB 579 PSIRILRGNTSIAVQRCSTWNRGQELTVESPVTSFTTNQSDLPFTFTQAOENLILAE 638
QY 648 SPASGEVYVDPKPELIPVNAIFREABEDLDVAKAVNGLFTSKCDALQTSVTVYVNOAN 707
DB 639 GVSTGEYFIDRIEIIIPVNPAREABEDLEAKKAVANLFTFRDGLQVNVTVYQVDOAN 698
QY 708 LVECLSDLYPNEKRLMDAVKAEKRLVQARNLQDTGPNRING--ENGWTSSTGTEVAB 765
DB 699 LVSCLSDEYGHDKKMLLEAVRAARLRSRERMLLQDPDNTTINSTEENGWKLNGVTISE 758
QY 766 GDVLFKORSRLTSAREIDTETPTLYLQOIDSLLKPYTRYKJLKGPIGSSQDLKILIR 825
DB 759 GGPFFKGRALQASAR----ENYPTIYOKVDASVLKPYTRYLQDGFVKSSQDLKILIR 814
QY 826 HRANQIVQVNDLPLPVLVNSCGGIDRCSEQQVYDANLALNENGNGNMSDSSHAFSP 885
DB 815 HKVHLVRNVPDNLVSDTYSDCSGSINRCDBQHQVDMQMDAHHHPMDCCRAAQTHFSS 874
QY 886 HDTGSEIDLNENTGIWVVKIPTPTNGYATLGNLELVEBGLSGETLERAQOQOQWQDKM 945
DB 875 YINTGDLANSVDQGIWVVKIPTPTNGYATLGNLELVEBGLSGETLERAQOQOQWQDKM 934
QY 946 AKRGASEKAYYAAQAIQDLRFADYQDQKLSGVMSDMLAQNLSQSIPIYVNDALPBI 1005
DB 935 GKRAEIDRVYLAQAQINHLFVDYQDQKLSGVMSDMLAQNLSQSIPIYVNDALPBI 994
QY 1006 PGNQTSPTLNRLOQAWNLVLDLNAIPNGDPNGSLSDNATSDVNVQOQSDTSVLVVP 1065
DB 995 PGINYEIYTELSDRLQOASLYTSRNVAVQNGDPNSGLDSMNWMTMDASVQOQGNMHPVLVS 1054
QY 1066 NWSVSOQFTVQPNRYVLRVYARKEGVGDGVYIIRDCANOTETLTFNACDDDTGVLSA 1125
DB 1055 HMDAQVSOQLRWPNCKTVLRVYARKEGVGDGVYIIRDCANOTETLTFNACDDDTGVLSA 1114
QY 1126 DQTSYITKTRFTPTQEWIDMSSTEGVFNIESVELLEE 1167
DB 1115 NDNSYITREVPYPTKIMWVEVSESGSFYIDSTIEFTQOE 1156

RESULT 15
US-09-589-477-72
Sequence 72, Application US/09589477
Patent No. 6570005
GENERAL INFORMATION:
APPLICANT: Schneckf, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth B.
APPLICANT: Walz, Michelle
APPLICANT: Stockhoff, Brian
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/589,477
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/886,615
FILING DATE: 1-JUL-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/674,002
FILING DATE: 1-JUL-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: MA-701C1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 1156 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-589-477-72

Query Match 51.3%; Score 3100.5; DB 4; Length 1156;
Best Local Similarity 52.0%; Pred. No. 8.5e-219;
Matches 615; Conservative 196; Mismatches 330; Indels 41; Gaps 13;
QY 1 MSNNONEVEILDASSSTVSVDNSVRYPLANDOTTTLQNNVYKDYLRMSGENPELFGNP 60
DB 1 MNNNONEVEILDAPHCQSPSDDDVRYPLASDPNAAQNNVYKDYQNTDEDTYDSYNP 60
QY 61 ETPISS-STVQNGIGIGVQVGLGALGVPFAGQIASFYFVIGOLMPSSTVSWEMMKQVE 119
DB 61 SLSISGRDAVOTALTVGVRILGALGVFPFSGQVSVFQFLNTLWPNVDTAIWEAFNRQVE 120
QY 120 DLDQKITSVRTALAGLQIGLGLDGLDYQKSLKNWLENDRDTRASVVTQVIALELDP 179
DB 121 ELVNOQITEPARNOALRLQGLGDSFNVYQSLQNLADLRDTRNLSSVRAQFIALDLDF 180
QY 180 VAKIPSPALSGQVPLLSVYAQAANLHLLLRDASIFGAEMGFTPGSEISTFYDQVTRTA 239
DB 181 VNAIPFVANGQVPLLSVYAQAANLHLLLRDASIFGEGNGFTQSEISYIYDRQLELTA 240
QY 240 QYSDYCVKMYNTGLDKGTNAASWLYKHQFRREMTLLVLDLVALFPNYDTRTYPIETTA 299
DB 241 KATNYCETWNTGLDRLRGNTESMLRYHQFRREMTLVLDLVALFPYDVRLYPTGSNP 300
QY 300 QLTREYVTDIVFNRETSGFCRRWSLNSDISPSEVASVIRSPHLFDILSEIEPYTTRA 359
DB 301 QLTREYVTDIVFNPPANVGLCRWGTNPYNTFSELENAPIRPPHLFDRLNSLTSSNR- 359
QY 360 GLPLNTEYLEYVYVGHISIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFOVRS 416
DB 360 -FPV-SNFMDSYSGHILRSYLYNDASVQSDSGLITTTTRATINPGVDGTR-----IEST 413
QY 417 GADLANVYAOVGVYVASFTL--LDKMTGSGVGGFTYSKPHMTMQVCTQNYNTIDBI 474
DB 414 AVDFRSALIGIYGNRASFPVPGGLFNGTTPSPANGG-----CRDLYDNDLELP 461
QY 475 ENPLRGVYS-HRLSHITSYSPKN-ASSPARVGNLPVPAWTRRSADVNTVYSDKITQI 532
DB 462 DE---STGSSTHRLSHVTFPFTQNAQGSIANAGSVPTVYVTRRDVDLNNITIPNRI 518
QY 533 PVVKANTLVSGTTVIKGPGTGNILKRTSSGPLAYTSVSVKSPLSQRYEARIRYASTTN 592
DB 519 PLVKASAPVSGITVLKPGPGTGGILRRITNGTFTGLRVTNPSLQRYRVRVPASSGN 578
QY 593 LRLPVTISGTRIYSINVNKTKMGDDLTNTFDLATIGT-----AFTFSNYSDSLTVGAD 647
DB 579 PSIRILRGNTSIAYQRCGSTMNGQELTYESFVTSBPTNQSDLPPTFTQAOENLILAE 638
QY 648 SPASGGEVYVDPKPELIPVNAITFRABEDLDVAKAVNGLFTSKKDALQTSVTDYQVNOAAN 707

DB 639 GVSTGSEYFIDRIEIIIPVNPAREABEDLEAAKAVANLFTTRDGLQVNVTDYQVQAAAN 698
QY 708 LVECLSDELYPNEKRLMDAVKEAKRLVQARNLQDTGFNRING--ENGWTSFGIEVAR 765
DB 699 LVSLSDSEQYGHDKWMLLEAVRAAKRLSRNRLQDPDFNTINSTENGWFKASNGVTISE 758
QY 766 GDVLFKORSRLTSAREIDTETTYTYLYQQIDSELLKPYTRYLKGPIGSSODLEIKLIR 825
DB 759 GGPFFKGRALQALASAR----ENTPTIYQKVDASVLPKPYTRYLDGPGVKSQDLEIDIH 814
QY 826 HRANOIVKVPNDNLLPDVLPVNSCGGIDRCSEQYVDANLALNNGNNGNMSDSSHAFSP 885
DB 815 HKRVHLVKNVPNDNLVSDTYSQSCSGINRDEQHOVDMLDAHHHPMDCCAAQTHEPSS 874
QY 886 HDTGEBIDNMENTGIWVVFKIPTNGYATIGNLELVEBGLSGETLBRRAQQEQOQMDNM 945
DB 875 YINTGDLNASVDQGIWVVLKVRITTDGYATLGNLELVEVGLSGESLREQRDNAMNAEL 934
QY 946 ARERGASEKAYYAKAIDELPADYDOKLNSGVENSMDLAAQNLVOSIPVYVNDALPEI 1005
DB 935 GRKRABIDRVYLAQAQAINHLFVDYQDQQLNPEIGLAINESNLVESISGVISDTLLQI 994
QY 1006 PGMNYSFTLTLNRLQAAWNLXDLRNALPNGDFRNLSDWNATSDVANVQQLSDTSVLVIP 1065
DB 995 PGINYIYIELSDRLQQAASLYTTSRAVQNGDFNSGLDSWNTTMDASVQDGNMHLVLS 1054
QY 1066 MNNQSQSQFTVQPNRYVYLVRTARKEGVGDGVYIIRDGANQTETLTNFCODDGTGVL 1125
DB 1055 HMDAQVSQQLRVNPNCKYVLRVLTARKVGGDGYVYIRGGAHQSTLTFLNACDYDNGTVV 1114
QY 1126 DQTSYITKTVEPTPSTEQWIDMSETEGVNIESVELVLEE 1167
DB 1115 NDNSYITEBVVFPYTKHMMVRSSESGSPYIDSIEPIBTQB 1156

Search completed: June 21, 2004, 13:40:18
Job time : 27 secs

Db 455 IDSGEELTYEN---YQSYSHRVSVYITSPKISTGTV--LGVVPFGWTHSSASRNNPIY 509
Qy 526 SDKITQIPVKAHTLVSGTTWIKGP-FTGNNILKRTSSGPLAY---TSVSVKSPLSORY 581
Db 510 ATKISQIPINKASRTSGAVNFOGLYNGGPMKLSGSGQVINLRVATDAKG-ASQRY 568
Qy 582 RARIRYASTTNLRLPVTTIS-----GTRIYSINVNKTWKGDDLTENTPDLATIGTA-P 633
Db 569 RIRIRYASDRAGKF--TISRSRPNPATYSIASIATNTMTNASTNLTSTFYAYSGPINL 626
Qy 634 TFSNYSDSLTVGADSFASGVBGVYDKFELIPVNATFEAEEDLDVAKKAVNGLFTSKDAL 693
Db 627 GISGSRTPDISITKEAGANLYIDRIEPIPVNTLFEAEEDLDVAKKAVNGLFTSKDAL 686
Qy 694 QTSVTDYQVQQAANLVECLSDLYPNKRMWDVAKEARLVQARNLLQDTGPNRNGEN 753
Db 687 QTSVTDYQVQQAANLVECLSDLYPNKRMWDVAKEARLVQARNLLQDTGPNRNGEN 746
Qy 754 GWTGSGTGEVAGDVLFKDRLSLTSAREIDTETPTVLYQOIDSLLKPTRYKLGFI 813
Db 747 GWTGSGTGEVAGDVLFKDRLSLTSAREIDTETPTVLYQOIDSLLKPTRYKLGFI 806
Qy 814 GSSQDLBIKLRHRANQIVKVPDNLPLDVLVNSCGGIDRCSEQQVYVDANLLENNGEN 873
Db 807 GSSQDLBIKLRHRANQIVKVPDNLPLDVLVNSCGGIDRCSEQQVYVDANLLENNGEN 866
Qy 874 GNMSSDSHAFSHIDTGEIDLNENTGIVVVPKIPPTNGYATLGNLELVBEGPLSGETLER 933
Db 867 GNMSSDSHAFSHIDTGEIDLNENTGIVVVPKIPPTNGYATLGNLELVBEGPLSGETLER 926
Qy 934 AQOQEQWQDMARKRGASEKAYAAQADRLPADYQDQKLSGVMSDMLAAQNLVQS 993
Db 927 AQOQEQWQDMARKRGASEKAYAAQADRLPADYQDQKLSGVMSDMLAAQNLVQS 986
Qy 994 IPYVNDALPEIPGMYNTSTFELNRLQQAANLVDLRNAIPNGDFRNLSDWNATSDVNV 1053
Db 987 IPYVNDALPEIPGMYNTSTFELNRLQQAANLVDLRNAIPNGDFRNLSDWNATSDVNV 1046
Qy 1054 QQLSDTSVLVIPNNNSOVQFTVQPNRYVRLVTRARKEGVDGVIIRDGANQTETLTF 1113
Db 1047 QQLSDTSVLVIPNNNSOVQFTVQPNRYVRLVTRARKEGVDGVIIRDGANQTETLTF 1106
Qy 1114 NICDDTGVLSADQTSYITKTVEPTPSTEQWIDMSETE 1152
Db 1107 NICDDTGVLSADQTSYITKTVEPTPSTEQWIDMSETE 1145

RESULT 7

US-08-789-449-2
; Sequence 2, Application US/08789449
; Patent No. 5824878
; GENERAL INFORMATION:
; APPLICANT: Ohba, Michio
; APPLICANT: Iwahana, Hidenori
; APPLICANT: Sato, Reichi
; APPLICANT: Suzuki, No. 5824878ukazu
; APPLICANT: Ogiwara, Katsutoshi
; APPLICANT: Sakanaka, Kazunobu
; APPLICANT: Hori, Hidetaki
; APPLICANT: Asano, Shouji
; APPLICANT: Kawasugi, Tadaaki
; TITLE OF INVENTION: No. 5824878el Microorganism and Insecticide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/789,449
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,887
; FILING DATE: 08-JUL-1994
; APPLICATION NUMBER: US 07/915,203
; FILING DATE: 23-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/K 301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-789-449-2

Query Match 60.6%; Score 3662.5; DB 2; Length 1149;
Best Local Similarity 63.8%; Pred. No. 4.6e-260;
Matches 752; Conservative 109; Mismatches 257; Indels 61; Gaps 21;

Qy 1 MSPNNQNEYETLDASSSTVSVDNSVRYPLANDOTTLLNMNMYKDYLRMSEGENPBLFGNP 60
Db 1 MSPNNQNEYETLDASSSTVSVDNSVRYPLANDOTTLLNMNMYKDYLRMSEGENPBLFGNP 60
Qy 61 ETPISS-STVGTGIGVGOVLGALGVPPAGQIASFYISVIGOLMPSSTSVVMEMIKQVE 119
Db 61 GTFISAQDAVGTGIDIVSTIISGLGIPVLGEVFSGLSLGLMFSNNENVMQIFPNRVE 120
Qy 120 DLIDOKITDSVRKTPALAGLQGLDGLDVYKSLKNLENRNDTRARSVVVTOYIALELDF 179
Db 121 ELIDOKILDSVRSRAIADLANSRITAVEVYQNALEDWRKNPHSTRSAALVKERFGNAEIL 180
Qy 180 VAKIPSPAISQOEVPILLSVYAAQANLHLLLDASIFGAEWGFTPGELSTFYDROVTRTA 239
Db 181 RTNMGSPSQTNVETPLPTPTAQAASLHLLVWRDVOIQYKGEWGPONDIDLFTYKQVSYTA 240
Qy 240 QYSDYCVKWNVTGLDKLGTNAASWLKYHOFRRMTLLVLDLVALFPNYDTTPTPIETTA 299
Db 241 RYSDHCQWYNAGLNKLGCTGAKQWVDYNRFRKNVNVLDLVALFPNYDARIYPLETNA 300
Qy 300 QLTRFVYTDPIVFNRETSGGFCRRWSLN-----SDI-----SFSVESAVIRSHPLFDIL 349
Db 301 ELTREIPTDPV-----GSYVTQSSSTLISWYDMPAALPSFSTLEN-LLRKPDFFTLL 352
Qy 350 SEIEFYTTTRAGLPLNNT-EYLEYVWGHISIKYKNTWASSALERNYCTIISNKKIYYDLANK 408
Db 353 QEIRMYTS---FRQNGTIEYNYMGQRLTUSIYVGSF--NKYSGLAGASDIIPVGON 407
Qy 409 DIFQVRSILGADLANVYAOVGVVPVASFLLLDKNTGSGSVGGFTYKSPKHTTMOVCTQNYNT 468
Db 408 DIYRV--VWTVIGRYTNSLLGVNPTP-YFSNNTOK-----TYSKP-----KQFAGGIKT 454
Qy 469 ID---EIPPENEPLSRGSHRLSHITSYSFKNASSPARYGNLPVANTHRADVTNTVY 525
Db 455 IDSGEELTYEN---YQSYSHRVSVYITSPKISTGTV--LGVVPFGWTHSSASRNNPIY 509
Qy 526 SDKITQIPVKAHTLVSGTTWIKGP-FTGNNILKRTSSGPLAY---TSVSVKSPLSORY 581
Db 510 ATKISQIPINKASRTSGAVNFOGLYNGGPMKLSGSGQVINLRVATDAKG-ASQRY 568
Qy 582 RARIRYASTTNLRLPVTTIS-----GTRIYSINVNKTWKGDDLTENTPDLATIGTA-P 633
Db 569 RIRIRYASDRAGKF--TISRSRPNPATYSIASIATNTMTNASTNLTSTFYAYSGPINL 626

QY 634 TFSNTSDSLTVGADSFASGGEVYVDFKELIPVNATFEAREEDLDVAKAVNGLFTSKDAL 693
Db 627 GISSRSRTDPSITTEAGAAANLYDIRIEIPVNTLFEAREEDLDVAKAVNGLFTSKDAL 686
QY 694 QTSVTDYQVQANLVECLSDLYPNEKRWLDVAKVRLVQARNLQDTCFNRLNGEN 753
Db 687 QTSVTDYQVQANLVECLSDLYPNEKRWLDVAKVRLVQARNLQDTCFNRLNGEN 746
QY 754 GWTGSTGIEVARGDVLFKDRSLRLTSAREIDTETPTLYLYQQIDSLKPYRYKLKGP 813
Db 747 GWTGSTGIEVARGDVLFKDRSLRLTSAREIDTETPTLYLYQQIDSLKPYRYKLKGP 806
QY 814 GSSQBLEIKLRHRANQIVKQVNDLPLVLPVNSCGGIDRCSBQYVYDANLALNNGEN 873
Db 807 GSSQBLEIKLRHRANQIVKQVNDLPLVLPVNSCGGIDRCSBQYVYDANLALNNGEN 866
QY 874 GNMSSDSHAPSFDITDGEIDMLNENTGIWVVKPIPTTNGVATLGNLVEBGPISGETLER 933
Db 867 GNMSSDSHAPSFDITDGEIDMLNENTGIWVVKPIPTTNGVATLGNLVEBGPISGETLER 926
QY 934 AQOQEQWQDKWARKGASEKAYYAAKQAIIDLFAFYQDQKLNKSGVMSDMLAAQNLVQS 993
Db 927 AQOQEQWQDKWARKGASEKAYYAAKQAIIDLFAFYQDQKLNKSGVMSDMLAAQNLVQS 986
QY 994 IPYVNDALPEIPGMNYSFTTELTLNLOQAWNLVLRNAPNGDPRNGLSNWNATSDVNV 1053
Db 987 IPYVNDALPEIPGMNYSFTTELTLNLOQAWNLVLRNAPNGDPRNGLSNWNATSDVNV 1046
QY 1054 QQLSDTSVLVLPNNNSQVSOQPTVQPNRYVLRVTRKGVGDGVYIIRDGANQETTLTF 1113
Db 1047 QQLSDTSVLVLPNNNSQVSOQPTVQPNRYVLRVTRKGVGDGVYIIRDGANQETTLTF 1106
QY 1114 NICDDDTGVLSDTQDSYITKTVETFTPTSTEQVWIDMSET 1152
Db 1107 NICDDDTGVLSDTQDSYITKTVETFTPTSTEQVWIDMSET 1145

RESULT 8
US-08-315-468-4
Sequence 4, Application US/08315468
Patent No. 5554534
GENERAL INFORMATION:
APPLICANT: Michaels, Tracy Ellis
APPLICANT: Foncecrada, Luis
APPLICANT: Narva, Kenneth E.
TITLE OF INVENTION: Process for Controlling Scarab Beetles
TITLE OF INVENTION: with Bacillus thuringiensis Isolates
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL USA
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/315,468
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/014,941
FILING DATE: 01 FEB 1993
APPLICATION NUMBER: 07/828,430
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/808,316
FILING DATE: 16-DEC-1991

ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MAY3.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1169 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: kumamotoensis
INDIVIDUAL ISOLATE: 50C
IMMEDIATE SOURCE:
LIBRARY: Lambdagem-11 library of L. Foncecrada
CLONE: 50C(b)
US-08-315-468-4

Query Match 52.2%; Score 3152; DB 1; Length 1169;
Best Local Similarity 54.0%; Pred. No. 1.4e-222;
Matches 653; Conservative 179; Mismatches 294; Indels 84; Gaps 25;

QY 1 MSPNNQNEVEIILDASSSTVSNSVRYPLANDQTTTLQNNYKDYLRMSGENPELFGNP 60
Db 1 MSPNNQNEVEIILDASSSTVSNSVRYPLANDQTTTLQNNYKDYLRMSGENPELFGNP 60
QY 61 ETFPIS-SSTVQTGIGVQVQALGALGVPPAGQIASFYSPVQGLWPSSTVSVMEMKQVE 119
Db 61 EYPLSEQDAVKAIDIVGKLLTGLGVPPVGVPIVSLYTQLIDILWPSKQKQSEIFMEQVE 120
QY 120 DLIQKITDSVRKTLAAGLQGLDGLDVTQKSLKNWLENRNDTRASVVTQVIALEDP 179
Db 121 ELINQKIAEYARNKALSELEGLNNYQLYLTALKEENKPNGRALRDVNRREILDSLF 180
QY 180 VAKIPSPAIJSGORVPLLSVYAOANLHLLLRDASIFGAEWGTGPGCIBISTFYDROVTRTA 239
Db 181 TQYMFSPRVTFEVPFLVTTWAANLHLLLRDASIFGAEWGTGPGCIBISTFYDROVTRTA 240
QY 240 QYSDYCVKYNQGLDKLKGTAASMLKYHQPREMTLLVLDLVALFPNVDTRYPIETTA 299
Db 241 EYSDHCVKYETGLAKLKGSSAKQWIDYNQPREMTLLVLDLVALFPNVDTRYPIETTA 300
QY 300 QLTREVTDPVFNRETSGPCRRWSLNSDISPSEVESAVIRSPHLFDILSEIFPTTRA 359
Db 301 QLTREVTDPVFNRETSGPCRRWSLNSDISPSEVESAVIRSPHLFDILSEIFPTTRA 356
QY 360 GLPLANTVLEYVWGHSIKYKNTNASSALERNYGT-----ITSNKIKYDLANKDIQVR 414
Db 357 S--FTSDRYPRYAGHQISYKHIGTSSTFTQYGTQNLQSTN----PFTNYDIYKTL 410
QY 415 SLGADLAN-----YTAQVYGVYPVASFLLDKNTGSGSVGGFTYSKPHHTVQVCTQNTY 469
Db 411 SNGAVLLDIVPGYVTPFGMPETEPFMVQNLNTRKT--LTY-KP-ASKDIDRTDSE 466
QY 470 DRIPPE--NEPLSRGYSHRLSHITSYFSKNASSPARYGNLVPFANTHSADVTNTVSD 527
Db 467 LELPPEPESQPNYSYSHRLGHIT--PIYSSSTST-----YVPVFSWTHRSADLNTVKS 520
QY 528 KITQIPVKAHTLVSGTIVKPGPTGNGILKRTSSGPLAYTSVSKSPISQRYRARIY 587
Db 521 EITQIPGKSSSTIGRNTYIIKRGYTGGLVALTDR--IGSCFQMFPEQSQRIRIY 578
QY 588 AS--TTNLELP-VTISGTRIYSINVKNT-MNKGD-DLTNTNFDLATIGTAFSTNSDSL 642
Db 579 ASNETSYISLYGLNQSQT----LKFQNTYSNKNENDLTND-----FKYIEYPRVI 625

Qy 643 TVGADS-----PASGGEVYV-DKELIPVNAATPEAREDDLVAKAVNGLTSSKDAL 693
Db 626 SVNASSMIQSLSIGITQNTNLLFIPVDETFEABTDLEAAKAVNALFTNTDGL 685
Qy 694 QTSVTDYQVNAANLVECLSDLEYFNKRLMDAVKAEKELVQARNLLQDTGFNRINGEN 753
Db 686 QPGVTDYEVNAANLVECLSDLEYFNKRLMDAVKAEKELVQARNLLQDTGFNRINGEN 745
Qy 754 GWTGSGTIEVAGVNLVKORSLRLTSAREIDTETPTLYLQOQDESLLKPYTRYKLKGP 813
Db 746 GWTASTGIEVIEGDAVFKGYRLPGAREIDTETPTLYLQOQDESLLKPYTRYKLKGP 805
Qy 814 GSSQDLKILRIPANQIVKRNPDNLDPVLVNSCGGIDRCSEQOQVVDANLALENNEN 873
Db 806 GSSQGLIYIRHQTNRIVKRNPPDLDLPVPPVNDGRINRCSQKYVNSKLEVEN--- 862
Qy 874 GNMSSDSHAFSFIHDTGEIDLENVTGIWVFKIPTTNGYATLGNLELVEGSPSGETLER 933
Db 863 ---SGRAHEFSIPIDTGEIDYENAGIIVGFKITDPGYATLGNLELVEGSPSGDALER 919
Qy 934 AQOQEQWQDMARKRGASKEYAQAQADRLFADYODQKLGNGVMSDMLAQNVLQS 993
Db 920 LQKEEQWQKQMTARRBETDRRYMASKQAVDRLYADYQDQQLNPNVEITDLTAQDLIQS 979
Qy 994 IPVYNDALPEICMNYTSTELTNRLOQANLYDLRNPNGDPFRNGLSDNATSDVNY 1053
Db 980 IPVYNEMFPEICMNYTKETLDRLOQANLYDLRNPNGDPFRNGLSDNATSDVNY 1039
Qy 1054 QQLSDTSVLVTPNNSOVSOQTPQPNRYVLRVYRKARKEGVDGWIIRDCANOTETLTF 1113
Db 1040 QQINHTSVLVPWNEQVSKFTVQPNQRYVLRVYRKARKEGVDGWIIRDCANOTETLTF 1099
Qy 1114 NICDDTGVLSADQTS---YIT-----KTVEFTPTSPTEQWIMDSMETEGVFI 1157
Db 1100 SASDYDTNGMYDQASNTNGYNTNSVYMIKPKRTVDISSVYVQWIBISETEGTFYI 1159
Qy 1158 ESVELVLEE 1167
Db 1160 ESVELVIDVE 1169

RESULT 9
US-08-532-547-5
Sequence 5, Application US/08532547
Patent No. 5861543
GENERAL INFORMATION:
APPLICANT: LAMBERT, BART
APPLICANT: JANSSENS, STEFAN
APPLICANT: VAN AUDENHOVE, KATRIEN
APPLICANT: PEPEROEN, MARNIX
APPLICANT: VAN RIE, JEROEN
TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND THEIR
TITLE OF INVENTION: INSECTICIDAL PROTEINS.
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532.547
FILING DATE: 06-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.

REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 2121-109P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-532-547-5

Query Match 51.7%, Score 3123, DB 2, Length 1157;
Best Local Similarity 52.3%, Pred. No. 1.9e-327;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

Qy 1 MSPNNQNEYELDASSSTSVSDNSVRYPLANDQTTTLQNNYKYDYLMSSEGENPELFGNP 60
Db 1 MNRNNQNEYELIDAPHCCFSDDDVRYPPLASDPNAALQNNYKYDYLQWDEDTDSYNP 60
Qy 61 ETFISS-STVQTGIGIVGVLGALGVPAGQIASFYSPFVQGLWPSSTSVVEMINKQVE 119
Db 61 SLSISGRDAVQATLTVGRILGALGVPSPGQIVSFYQFLANTLMPVNDTAIWEAFMRQVE 120
Qy 120 DLIDQKITDSVRKTAGLAGLQGLDGLVYQSKLNNLENRNDTRARSVVVQYIALELDF 179
Db 121 ELVNNQITFEPARNQALRLOGLGDSFNVYQSLQNLADRNDTRNLSSVRAQFIADLDF 180
Qy 180 VAKTIPSAISOEVPILSVVAQAANLHLLLRDASIFCAEWGFTPGELSTPYDROVTRTA 239
Db 181 VNAIPLFAVNGQQVPLSVVAQAANLHLLLRDASLFGEGWGTGELSTPYDROBELTA 240
Qy 240 QYSDYCVKYNATGDKLKTNAASWLKTHQFRREMTLLVLALFPNYDTTYPIETTA 299
Db 241 KYTNYCETWNTGLDRGNTSMLRYHQFRREMTLLVLVAVLFPYDRLYPTGSNP 300
Qy 300 QLTREVTDPVFNRETSGGFCRAWLSNDISFSEVESAVIRSPLFDLSIBFYTTTRA 359
Db 301 QLTREVTDPVFNPPANVGLCRWGTNPYNTFSELENAFIRPPHLPRLNLSLTSSNR- 359
Qy 360 GLPLNNTLEYVWVGHSLIKYKNTNASSALERNYGTITSNKIKY---YDLANKDIFQVRS 416
Db 360 -FPV-SSNFMDYWSGHTLRRSYLNDASVQESYGLITTRATINFGVGTNR----IEST 413
Qy 417 GADLANYAQAIVGVYASFTL--LDKNTGSGSVGGFTYSKPHTTMQVCTQNTYNTIDEIPP 474
Db 414 AVDFRSALIGIYGVNRASFVPGLFNGTTSFANGG-----CRDLVDNDLELP 461
Qy 475 ENEPRLSGYS-HRLSHITSYSFSKN-ASSPARYGMLPVFATWHRSAADVNTYVSDKITQI 532
Db 462 DE---STGSSTHRLSHVTFPSQTNQAGSIANAGSVPTVYVMTTRRDVLDLNTIPNRITQ 518
Qy 533 PVKANTLVSGTIVIKGPGFTGMLKETSGLAYTSVSVKSPLSQSORARIRVASTTN 592
Db 519 PLVKASAPVSGTIVIKGPGFTGGGLRRTTNGTFCGLRVTVVNSPLTQQRLRVRFSTGN 578
Qy 593 LRLFTVTSITRIYSINVNKTKNGDGLTFNFP---DLATIG---TAFTPSNYSDSLTVGA 646
Db 579 FSIRVLRCGVSIGDVLGSGTWNRGELTYESFPTTGTGFPFPPTFQAQELTVNA 638
Qy 647 DSPAGSGEVYVDKPELIPVNATFEAREDDLVAKAVNGLTSSKDALQTSVTDYQVNOQA 706
Db 639 EGVSTGGEYIDRIEIVPNPAREAREDELEAKAVASLFTTRDGLQVNVTDYQVDOQA 698
Qy 707 NLVCLSDLEYPNBCKMLDVAKEKRLVQARNLLQDTGFNRING--ENGWGTGSGTIEVA 764
Db 699 NLVCLSDLEYPNBCKMLDVAKEKRLVQARNLLQDTGFNRING--ENGWGTGSGTIEVA 758
Qy 765 EGDVLPKDRSLRLTSAREIDTETPTLYLQOQDESLLKPYTRYKLKGFSSQDLLEIKLI 824
Db 759 EGGPFPKRALQLASAR---ENYPTIYQKVDASVLPKPYTRYLDFGVKSSQDLLEIKLI 814

```
QY 825 RHRANQIVKYNVDPNLLPDVLPVNSCGGIDRCSEQQYVDANLALNNGENGMSSDSHAPS 884
DB 815 HHHKVLKYNVDPNLLPDVLPVNSCGGIDRCSEQQYVDANLALNNGENGMSSDSHAPS 874
QY 885 PHIDTGEIDLNENTGKIWWVFKLPTTNGVATLGNLELVEBGLSGTTLERAAQOQOQWQDK 944
DB 875 SYINTGDLNASVDQGIWVVKVTRTDTGATLGNLELVEBGLSGTTLERAAQOQOQWQDK 934
QY 945 MARKGASKEYAAQADRLFADYDQOKLNSGVMSDMLAAQNLVQSIPIVYNDALPE 1004
DB 935 LGRKAEIDRVTLAAQAINHLFVDYDQOKLNSGVMSDMLAAQNLVQSIPIVYNDALPE 994
QY 1005 IFCMYTSTFTELTLNRLQANLYDLNAPLNGDFNGLSDNATSDVAVQQLSDTSVLVI 1064
DB 995 IFCMYTSTFTELTLNRLQANLYDLNAPLNGDFNGLSDNATSDVAVQQLSDTSVLVI 1054
QY 1065 PWNNSQVSOQPTQPNRYVLRVARTARKGVBGGYVIRIDGANTOTETLTFNICDDDTGVLS 1124
DB 1055 SHWDAQVSOQLRVNPNCKYVLRVARTARKGVBGGYVIRIDGANTOTETLTFNICDDDTGVLS 1114
QY 1125 ADQTSYITKTVETFTSTEQWIDMSSETGVFNIESVRLVLEEB 1167
DB 1115 VNDNSYITEEVVFPETKHMVVESESGSFYIDSIETIETQE 1157
```

RESULT 10

```
US-08-379-656B-5
; Sequence 5, Application US/08379656B
; Patent No. 5985571
; GENERAL INFORMATION:
; APPLICANT: Lambert Bart
; APPLICANT: Janssens, Stefan
; APPLICANT: Van Audenhove, Katrien
; APPLICANT: Peferoen, Marlinx
; TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND
; TITLE OF INVENTION: THEIR INSECTICIDAL PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESS: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,656B
; FILING DATE: 23-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/01820
; FILING DATE: 12-JULY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 93400949.9
; FILING DATE: 09-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92402358.8
; FILING DATE: 27-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 2121-104P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEFAX: 248345
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-379-656B-5
```

```
Query Match 51.7%; Score 3123; DB 2; Length 1157;
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;
```

```
QY 1 MSPNNQNEIYELDASSSTSVSDNSVRYPLANDQTTTLQNMNKKDYLRMSEGENPELFCNP 60
DB 1 MNRNQNNEIYELDAPHCPCSDDDVRYPLASDPNAAQNMNKKDYLRMSEGENPELFCNP 60
QY 61 ETFISS-STVGTGIGVQVIGALGVPPAGQIATSPYIVGQVIGQVIGQVIGQVIGQVIGQV 119
DB 61 SLSISGRDAVQTLTVGRIILGALGVPPSGQIVSPYQPLNLTPVNDATIAEAFNRQVE 120
QY 120 DLIDQIKTDSVRKTAALAGLQGLGDVYVQSLKQWLENRNDTRARSVVVTVYALDELDF 179
DB 121 ELVNOQITEFARNQALARLQGLGDSFNYYQSLQNLADRNDRNLVVVRAQFIALDLDF 180
QY 180 VAKIPSPALSGOEVPLLSVYAAQANLHLLLDASIPGAEWGFTEGISTFTDQVTRTA 239
DB 181 VNAIPFPAVNGQVPLLSVYAAQANLHLLLDASIPGAEWGFTEGISTFTDQVTRTA 240
QY 240 QYSDYCVKMYNTGLDKAGTNAASWLKYHOPREMTLLVLDLVALPMPYDTRTYPIETTA 299
DB 241 KYTNYCETWYNTGLDRLRGNTESWLRYHOPREMTLLVLDLVALPMPYDTRTYPIETTA 300
QY 300 QLTREVVYTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHILFSLSEIFYTTRA 359
DB 301 QLTREVVYTDPIVFNRETSGGFCRRWSLNSDISFSEVESAVIRSPHILFSLSEIFYTTRA 359
QY 360 GLPLNNTLEYLVWGHSLKYKNTNASSALERNYGTITSNKIKY----YDLANKDIFQVRS 416
DB 360 -FPV--SSNFMQYMSGHTLRRSYLNDASVQEDSYGLITTRATINPGVDCGTNR----IBST 413
QY 417 GADLANYTAAQVYGVYASFTL--LDKNTGSGVGGFTYSKPHITTMQVCTQNTYNTIDSI 474
DB 414 AVDFRSALIGYVGNRASFPVGLFNGTSPANGG-----CRDLTDNDELPP 461
QY 475 ENEPLSRGYS--HRLSHITSYSPSKN--ASSPARYGNLVPFAMTHRSADVTNTVYSKITQI 532
DB 462 DE---STGSSHTLSHVTFPSPQTNQAGSIANAGSVPTVVTTRDVLANTITPNTITOL 518
QY 533 PVKHAHTLVSGTIVKGPFGTGNLTKTSSGLAYTSVSVKSPISQRTAARIRYASTTN 592
DB 519 PLVKASAPVSGTIVKGPFGTGNLTKTSSGLAYTSVSVKSPISQRTAARIRYASTTN 578
QY 593 LRLPVTISGTIYSINVAKTNKGDILFNTF--DLATIG--TAFTFSNYSDSLTVGA 646
DB 579 PSIRVLRGGVSGIDVRLGSTMNRGQELIYSEFTREFTTGGPFPFPFTTQAEILTVNA 638
QY 647 DSFASGGEVYVDKPELIPVNAITFEAEEDLDVAKKAVNGLFTSKDALQTSVTDYQVNOAA 706
DB 639 EGVSTGGEYIIDRIEIVPNPAREAEEDLEAAKAVASLFTTRDGLQVNVTDYQVDOAA 698
QY 707 NLVECLSDLEYPNEKRLMADVAKELVQARNLLQDTGFNRING--ENGWGTSGTIEVA 764
DB 699 NLVSLCLSEQYGHDKKMLLEAVRAAKRLSRERLQDPDFTNTINSTEENGKASNGVTIS 758
QY 765 EGDVLFKDRSLRSLTSARBIDTETYPYVQQLDESLKPYTRYKLGKFGSSQDLRIKLI 824
DB 759 EGGPFYKGRALQALASR-----ENYPTIYQKVDAVSLKPYTRYLKGFKVSSQDLRIKLI 814
QY 825 RHRANQIVKYNVDPNLLPDVLPVNSCGGIDRCSEQQYVDANLALNNGENGMSSDSHAPS 884
DB 815 HHHKVLKYNVDPNLLPDVLPVNSCGGIDRCSEQQYVDANLALNNGENGMSSDSHAPS 874
QY 885 PHIDTGEIDLNENTGKIWWVFKLPTTNGVATLGNLELVEBGLSGTTLERAAQOQOQWQDK 944
DB 875 SYINTGDLNASVDQGIWVVKVTRTDTGATLGNLELVEBGLSGTTLERAAQOQOQWQDK 934
```


APPLICANT: VAN AUDENHOVE, KATRIEN
APPLICANT: PEPEROEN, MARNIX
APPLICANT: VAN RIE, JEROEN
APPLICANT: VAN AARSEN, ROEL
TITLE OF INVENTION: NEW BACILLUS THURINGIENSIS STRAINS AND THEIR
TITLE OF INVENTION: INSECTICIDAL PROTEINS.
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA: #1.0, Version #1.25 (EPO)
APPLICATION NUMBER: US/09/471,177
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: US/09/019,809
FILING DATE: 02-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 2121-135P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-471-177-5

Query Match 51.7%; Score 3123; DB 4; Length 1157;
Best Local Similarity 52.3%; Pred. No. 1.9e-220;
Matches 619; Conservative 195; Mismatches 327; Indels 42; Gaps 14;

Qy 1 MSPNNQNEYETLDASSSTVSVDNSVRYPLANDQTTTLQNNYKDYLRMSGEENPELFCNP 60
Db 1 MRRNNQNEYETLDAPHCCPSDDVRYPLASDPNAALQNNYKDYLTQWDBEDYDSINP 60

Qy 61 ETPISS-STVQTGIGVQVGLGALGVPFAGQIASFYFVQGLWPSSTVSVMIMQVZE 119
Db 61 SLSISGRDAVQTALTVGRILGALGVPPSGQIVSFYQPLTLNLPVNDTALWEAFMRQVE 120

Qy 120 DLIQDKITDSVRKTAGLQGLDGLDVYQSLKNWLENRNDTRARSVVVYQYIALELDF 179
Db 121 ELVNNQITTEPARNQALRQGLGDSFNYYQSLQNLADNRDNLNLSVRAQFIALDLDF 180

Qy 180 VAKIPSPAISQBYPLLSVYAQAANLHLLLRDASIFGAEGWFTPGESTIDYDQVTRTA 239
Db 181 VNAIPFVAVQVPLLSVYAQAANLHLLLRDASIFGAEGWFTPGESTIDYDQVTRTA 240

Qy 240 QYSDYCVKWNVTGLDKLKGNTNAASWLYKHQFRRENTLLVLDLVALPNYDTRTPIETTA 299
Db 241 KYTNVCTWNTGLDRLRGNTTSMWLRVHQFRRENTLLVLDLVALFPYDVRVLTGSNP 300

Qy 300 QLTREVTDPVFNRETSGGCRWNSLSDISPEVESAVIRSPLDILSEIFPYTTRA 359
Db 301 QLTREVTDPVFNRETSGGCRWNSLSDISPEVESAVIRSPLDILSEIFPYTTRA 359

Qy 360 GLPLNNTLEYVWCHSFKYKNTNASSALERNYGTITSNKIKY---VDLANKDIFQVRS 416
Db 360 -FPV-SSNFMDWSGHLRRSYLNDASVQEDSYGLIITTRATINPGVDGTR-----TEST 413

Qy 417 GADLANYYAQTGYVPYASFTL--LDXNTGSSVGGFTYSKPHTTMQVCTQWNTIDRIPP 474
Db 414 AVDFRSALIGYGVNRASFVPGGLFNGTTSFANGG-----CRDLVDTNDELPP 461

Qy 475 ENBPLSRGYS-HRLSHITSYSFSKN-ASSPARYGNLPVPAWTHRSADVNTVYSDKITQI 532
Db 462 DE---STGSSTHRLSHVTFFSFQTNQAGSIANAGSVPTVYVTRRDVLDLNTTTPNRIQL 518

Qy 533 PVKRAHTLVGTTVKGPGFTGGNILKRTSGPLAITSVSVKSPULSQRYRARIRVASTTN 592
Db 519 PLVKASAPVSGTTLVKGPGFTGGGLRRTTNGTFTGLVTVNSPLTQOYRLVREFASTGN 578

Qy 593 LRLFTTISGTRIYSINVNKTNKGGDLTFNPF---DLATIG---TAPFNSVDSITVCA 646
Db 579 PSIRVLRGVSIGDVRLGSTNMRGOELTYBSFTRFTTTPGFPNPPFTTQAEILTVNA 638

Qy 647 DSPASGGGVYVDKPELIPVNATFEAREEDLDVAKAVNGLFTSKKDALQTSVTDYVNOAA 706
Db 639 EGVSTGGGYIDRIEIVPVNPEAREAEDELEAKKAVASLPTTRDGLQVNVTDYQVQAA 698

Qy 707 NLVECLSDLEYPNEKRMMDAVKAKRLVQARNLLQDTGPNRING--ENGWTSGTGIEVA 764
Db 699 NLVSLSDBOYGHDKQMLLEAVRAAKRLSRERNLLQDPDFTNTINSTENGWKAASGVTTIS 758

Qy 765 EGDVLFYDRSLRLTSAREIDTETVYTYLQQIDSLKPKYTRYKLKGFSSQDLLEIKLI 824
Db 759 EGGPFFKGRALQLASAR-----ENPTYIYQKVDASVLKPYTRYLRDGFVKSSQDLLEIKLI 814

Qy 825 RHRANQIVKQVNPMLPDLVPVNSCGGIDRCSEQQYVDANLALNNGNNGNMSDPSHAPS 884
Db 815 HHKVKHLVKNVNDLVSDTYSGSCGSLNRCDEHQHVDMDQDAEHHPMDCCAAQTHBFS 874

Qy 885 PHIDTGEIDLNENGTIWWVFKIPTTNGVATIGNLELVEBGLSGTSLERAQOQOQWQDK 944
Db 875 SYINTGDLNASVDQGIWVLKVRITTDGYATLGNLELVEBGLSGSLEREQRDNKWNAB 934

Qy 945 MARYGASERKAYAAKQAIDELPAQYQDOKLNSGVEMSDMLAAONLVOSIPVYNDALPE 1004
Db 935 LGRKRAEIDRVYLAARQAINELFVDYQDQQLNPEIGLABINEASNLVSSISGVYSDTLIQ 994

Qy 1005 IPGNVYTSFTLNTLRQAQWNLVYLRNAINPMDGFRNGLSMDNATSDVNVVQLSDTSVLVI 1064
Db 995 IPGNIYEIYELSDRLQQAASLYTSRAVQNGDFNSGLDSNNTTMDASVQDQGNMFLVL 1054

Qy 1065 PNWNSQVSQPTVPNTVRYLVRTARKEGVGDGYVIIRDGANQTETLTFNICDDDTGVLS 1124
Db 1055 SHWDAQVSQQLRVNPNCKYVLRVTARKVGGDGYVTIRDAHQBHTLTFNACDYVNGTY 1114

Qy 1125 ADQTSYTKTVRFTPTSTEQWIDMSRTEGVNIESVELVLEEE 1167
Db 1115 VNDNSYITEEVVFPYPTKHMVVEVSESEGSFYIDSIFETQOE 1157

RESULT 14

US-09-002-285-72
Sequence 72, Application US/09002285
Patent No. 6369213

GENERAL INFORMATION:
APPLICANT: Schnepf, H. Ernest
APPLICANT: Wicker, Carol
APPLICANT: Narva, Kenneth E.
APPLICANT: Walz, Michelle
APPLICANT: Stuckhoff, Brian
APPLICANT: Muller-Cohn, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA

